THE UNIVERSITY OF ALBERTA

GEOSTATISTICS OF PETROLEUM RESERVES

by

MAURICE HONORE

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

IN STATISTICS

DEPARTMENT OF MATHEMATICAL SCIENCES

EDMONTON, ALBERTA

FALL 1998

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-34376-6

Canada

University of Alberta

Library Release Form

Name of Author: Maurice Honore

Title of Thesis: Geostatistics of Petroleum Reserves

Degree: Master of Science

Year this Degree Granted: 1998

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

12204 - 39 Avenue
Edmonton, AB
T6J 0N2

Date: October 2, 1998

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommended to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled "Geostatistics of Petroleum Reserves" submitted by Mr. Maurice Honore in partial fulfillment of the requirements for the degree of Master of Science in Statistics.

Doug Kelker

Dr. Doug Kelker (Supervisor)

Dr. Sam Shen

Dr. Clayton Deutsch

Date: October 2, 1998

ABSTRACT

This thesis discusses the concepts of geostatistics of petroleum reserves. In order to predict the amount of petroleum in a certain area, one must first record the amount of oil at selected well sites in the area. Then one models the variation in the data using a tool known as the semivariogram. Once a semivariogram model has been chosen, then by means of a method known as kriging, the average amount of oil in the area is predicted. If circumstances exist where kriging is not feasible, then alternative methods of prediction such as random kriging or regularization are used. If the intention is to study the dispersion characteristics of the data, then a method known as conditional simulations is used. Throughout this thesis, some of these methods are slightly modified in order to allow greater freedom in fitting the models to the data. As well, some of these methods are applied to real world examples.

TABLE OF CONTENTS

SECTION	PAGE
1. INTRODUCTION	1
2. BASIC GEOSTATISTICAL CONCEPTS	4
3. STATIONARITY	5
4. SEMIVARIOGRAM MODELS	7
A. The exponential model	7
B. The spherical model	11
C. Nugget Effect	14
D. Example (A): Fitting an appropriate semivariogram	16
5. NESTED MODELS	19
A. BACKGROUND	19
B. EXAMPLE (B)	19
6. ANISOTROPIES	21
A. Geometric Anisotropy	21
B. Zonal Anisotropy	24
7. HOLE EFFECTS	24
8. MODELLING AS A TIME SERIES	27
A. Background	27
B. Case(1): ARMA(1,0) with $\psi = 0.8$	31
C. Case(2): ARMA(1,0) with $\psi = 0.9$	31

D. Case(3): ARMA(0,1) with $\tau = 0.5$ 32

E. Case(4): ARMA(1,1) with $\tau = 0.3$ and $\psi = 0.9$ 32

F. Case(5): ARMA(1,1) with $\psi = 0.9$ and $\tau = 0.6$ 33

G. Case(6): ARMA(2,0) with $\psi_1 = 0.9$ and $\psi_2 = -0.05$ 33

H. Case(7): ARMA(2,0) with $\psi_1 = 1.8$ and $\psi_2 = -0.9$ 34

9. EXAMPLE (C) 35

10. PROPORTIONAL EFFECT AND QUASI-STATIONARITY 39

A. Proportional Effect 40

B. Cases 40

11. AUXILIARY FUNCTIONS 41

A. Mathematical Representation 42

B. Generalized Exponential Model 43

C. Generalized Spherical Model 45

12. KRIGING 46

A. Background 46

B. Kriging System 48

C. Matrix Form 49

D. Kriged Estimates 51

E. Influence of a Nugget Effect 52

13. MULTIVARIATE CASE 54

A. Coregionalization 54

B. Cokriging Equations 56

C. Data Reduction in Cokriging 59

14. EXAMPLE(D) 62

15. DRIFT 65

A. Background 65

B. Types of Drift 66

16. NONLINEAR GEOSTATISTICS 66

17. VECTORIAL SPACE PROJECTION THEOREM 67

18. LINEAR KRIGING PROCESSES 69

A. Background and Conditions 69

B. Cases 70

19. ESTIMATION VARIANCE 73

20. RANDOM KRIGING 75

21. DISPERSION VARIANCE 79

22. CONDITIONAL SIMULATIONS 80

23. REGULARIZATION 82

24. CONCLUSION 87

25. APPENDIX A: DATA FOR FITTING BASIC SEMIVARIOGAMS 88

26. APPENDIX B: DATA FOR FITTING A DOUBLE SPHERICAL MODEL 90

27. APPENDIX C: DATA FOR FITTING AN ANISOTROPIC MODEL 92

28. APPENDIX D: DATA FOR THE KRIGING AND AUXILIARY FUNCTIONS EXAMPLE 99

29. BIBLIOGRAPHY 102

SECTION 1 INTRODUCTION

The purpose of this thesis is to study the concepts of geostatistics, as applied to petroleum reserves. The bulk of the material used in this thesis came from the references Cressie, Hohn, Journel & Huijbregts, and Rendu. As well, the FORTRAN program in Deutsch was used to perform an estimation problem using the techniques laid out in these references. Geostatistics is a technique that predicts the existence and quantities of a variable (such as iron, copper) in space. This technique is used because after performing seismic tests to prove the possible existence of the variable in the area, a company wants to be able to predict the total quantity of the variable in the area. Before one can perform these techniques, data on the variable under study must be collected. The first step is to set up a coordinate system with the easting and northing directions as coordinates for each point in the area. The next step is to drill holes at selected points on the grid, and from the drill cores, determine the potential quantities of the variable at each point. Once this is done, an estimation procedure known as kriging is used to predict the quantities of the variable at other points on the grid, and eventually the total quantity of this variable. With regard to the estimation of petroleum reserves, an area is chosen after seismic tests prove the potential existence of petroleum in that area. A grid

is then set up for this area. At selected points on this grid, a certain geological procedure is performed in order to give us the initial potential of these reserves. These initial potential values are recorded. Kriging is then used to predict average initial potential of the area of interest, and the total initial potential in the area.

Some of the standard geostatistical formulas in the textbooks will be modified in a way that will allow greater freedom in fitting models to data. Some of these modifications will be subject to certain mathematical conditions that are necessary for a model to hold. The concept of stationarity, which is defined in terms of first and second moments, will be examined first, since most geostatistical models are based on this assumption. Kriging uses the spatial correlations between the samples to estimate the regionalized variable at the unknown points on the grid. A geostatistical method of describing this spatial variation is the semivariogram. Two basic types of semivariogram models, exponential and spherical, will be examined in great detail with the aid of several graphs. These concepts will then be applied to real world examples using either one of the basic models or a combination of the two. The notion of an anisotropy will then be considered, in order to better model the variation by considering direction as well as distance. Since the variation may sometimes appear somewhat periodic, the notion of a hole effect will then be looked at.

Once the semivariogram models have been discussed, then

another way of modelling the semivariograms will be presented in the form of a time series. The basic concepts of a time series will first be explained, then it will be shown how they relate to the semivariogram models. The time series models will then be graphed in order to show the similarities between the two ways of modelling.

The assumption of stationarity may not hold, so before considering the estimation procedure, the concepts of quasi-stationarity and drift will be looked at. Kriging depends on the semivariogram and will be discussed next. The estimates are obtained by solving a system of linear equations that minimize the estimation variance. The concept of cokriging will then be analysed to deal with the case where we are estimating the initial potential in an area using other variables in addition to the initial potentials at selected points in the area.

Alternative forms of estimation will then be looked at in the event that kriging may not be possible. If for example, the locations of the points are unknown, but randomly scattered, or if for some reason we cannot measure the initial potential value at certain points in the area of study, then methods known as random kriging and regularization are needed and will be examined.

SECTION 2 BASIC GEOSTATISTICAL CONCEPTS

Geostatistics is based on the concept of a regionalized variable(Hohn, pp. 2). A regionalized variable is a variable that is distributed in space. The purpose of geostatistics is to predict the values of this regionalized variable. In order to do this, one must first model the spatial variation of this variable. This can be accomplished by the use of a semivariogram (often denoted as $\gamma(h)$, where h is the distance between two points). The semivariogram is a graphical device that is used to model spatial continuity and autocorrelation. In estimating petroleum reserves, we attempt to measure and predict the initial potential(often denoted as IP) of the reserves. The IP is determined when geologists perform a DST(drill standard test). In doing so they drill a hole at a potential well site and they measure the water pressure and other variables of interest and from these they are able to determine the IP of this potential well site.

Let $z(x)$ = IP at well site x
and $z(x+h)$= IP at well site x+h.
The usual estimate of $\gamma(h)$ is

$$\gamma^*(h) = \Sigma_i [z_i(x)-z_i(x+h)]^2/2n \quad , \; i=1,2,.....,n$$

where n = the number of pairs of points on the grid that are h units apart.

The semivariogram $\gamma(h)$ is a measure of covariance and is similar to the familiar covariance of basic statistics. $\gamma^*(h)$, the experimental semivariogram is calculated using all pairs of wells separated by a distance h. This is done for all values of h up to as big as h can be such that there exists sufficient data to calculate the experimental semivariogram. Once this is done, these values are plotted on a graph with distance on the horizontal axis and $\gamma(h)$ on the vertical axis. For small h, the IP are about equal, i.e. as h goes to 0, so does $\gamma(h)$. For large h, the values become more independent, i.e. $\gamma(h)$ tends to level off at a certain h value. This value of h is called the range of the semivariogram. The range is simply the distance at which we see a transition from spatial correlation to the absence of spatial correlation. The corresponding $\gamma(h)$ is called the sill.

SECTION 3 STATIONARITY

Stationarity is an assumption that must hold in order for the semivariogram models to be valid. It assumes that the first and second moments of the observed random function remain invariant across the area being studied. For the regional variable z(x)

at point x, the distribution function has the following property: $E[z(x)] = \mu(x)$, and this value can depend on the location x. This is our first order moment. The following properties must hold in order for first order strict stationarity to exist(Journel & Huijbregts, pp. 30-32):

$E[z(x_0)-z(x_0+h)] = \mu(h)$, where $\mu(h)$ is a function only of the distance h, and it is independent of the point x_0. i.e. the semivariogram is independent of the point x for all possible distances h on the grid being studied.

The two conditions for second order stationarity are:

(1) $E[z(x)] = \mu(x) = \mu$, the expectation of z(x) is independent of x, i.e. it is invariant across the grid being studied

(2) The covariance for each pair of regionalized variables $\{z(x), z(x+h)\}$depends only on the separation distance h, with covariance $C(h) = E\{z(x+h)z(x)\} - \mu^2$, for all x. If the covariance is stationary, the variance and semivariogram are defined to be stationary, thus

$$C(0)=E[(z(x)- \mu)^2]= Var(z(x))$$

$$\gamma(h)=E[(z(x+h)-z(x))^2]/2 = C(0)- C(h)$$

At this point, it is important to note that there is a restriction on the choice of semivariogram due to the one to one correspondence between the semivariogram and the covariance, which must be positive definite.

SECTION 4 SEMIVARIOGRAM MODELS

The two types of semivariogram models that will be analysed are the (1) Exponential Model and (2) Spherical Model(Hohn, pp. 25-28).

A. EXPONENTIAL MODEL

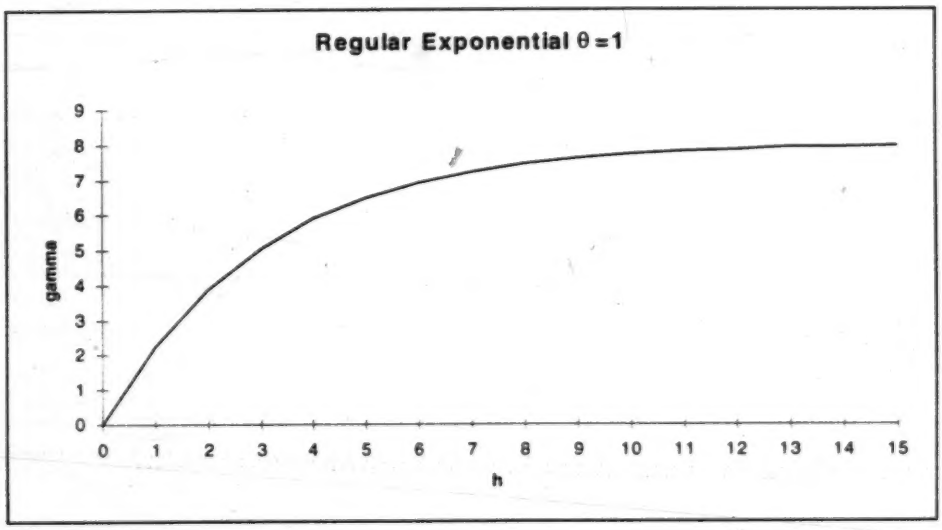
The basic form of the exponential model is $\gamma(h)=C[1-\exp(-(h/a))]$(Hohn, pp. 27), but in order to allow more freedom in choosing the parameters and to give a better fit of the model to the data, this model has been modified and is now the following more general model:

$$\gamma(h)=C[1-\exp(-(h/a)^{\theta})], \text{ for } 0<\theta<2,$$

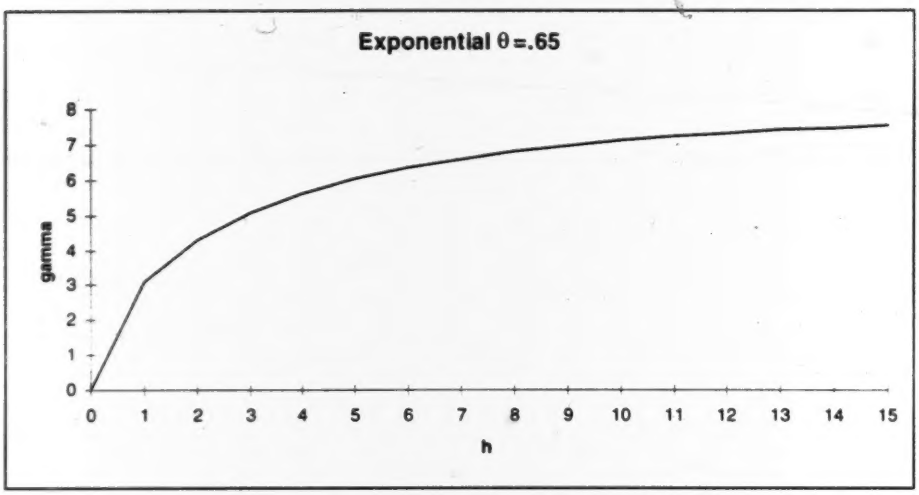
to ensure positive definiteness, where the C and a are the sill and range, respectively, as stated earlier. The θ value is a parameter that is chosen in order to give us the best fit. Finding the values of C, θ, and a is done using a trial and error process. We first look at the graph of the plot of the experimental semivariogram and choose C and a based on the pattern of the experimental semivariogram. For instance, if a levelling of $\gamma(h)$ occurs at the point $(h_0,\gamma(h_0))$, $a=h_0$ and $C=\gamma(h_0)$ would be chosen as the range and sill respectively. After having chosen these values, θ is then chosen in order to provide

the best possible fit to the data. A special case of the exponential model is the case where $\theta=2$, this is known as the Gaussian model. At this point it is important to note that $\gamma(h)$ approaches the sill-value, C, asymptotically.

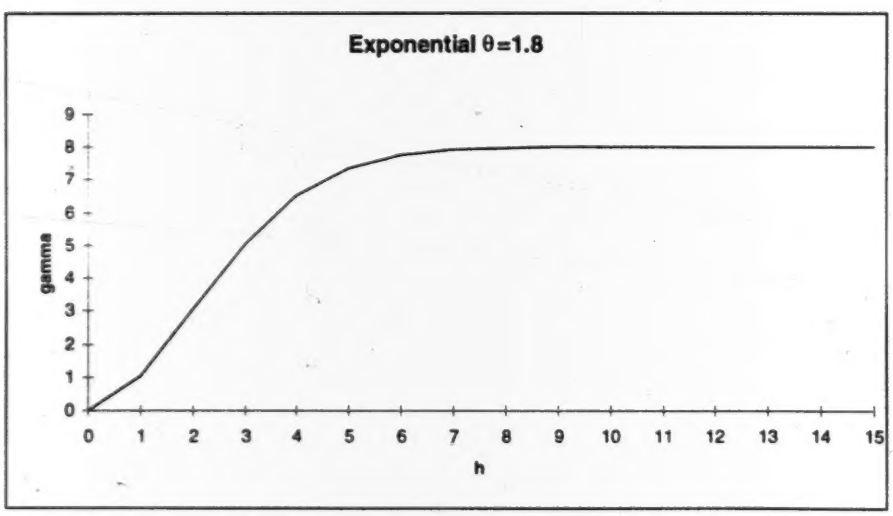
Case (1) Regular Exponential with $\theta = 1$:



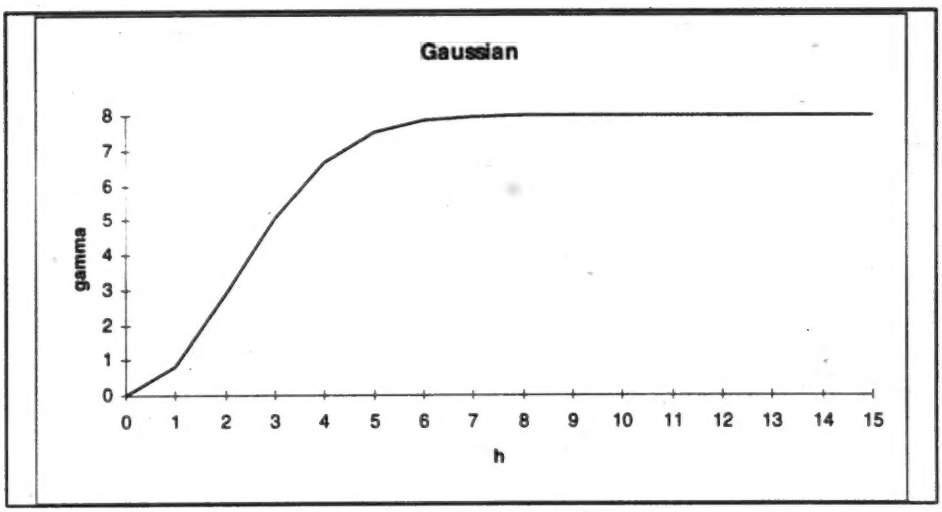
Case(2) Exponential with θ = 0.65(more hump-shaped):



Case(3) Exponential with θ = 1.8(slightly more parabolic shaped):



Case(4) Gaussian(Exponential with $\theta = 2$):



B. SPHERICAL MODEL

The basic form of the spherical model is

$\gamma(h) = C[(3/2)(h/a)-(\frac{1}{2})(h/a)^3]$ (Hohn, pp. 26).

But as in the case of the exponential model, to allow more freedom in choosing the parameters and to give a better fit of the model to the data, the model has been modified to give the following general model:

$$\gamma(h) = \begin{cases} C\{(\phi/(\phi-1))*(h/a) - (1/(\phi-1))(h/a)^c\} & , h \leq a \\ C & , h>a \end{cases}$$

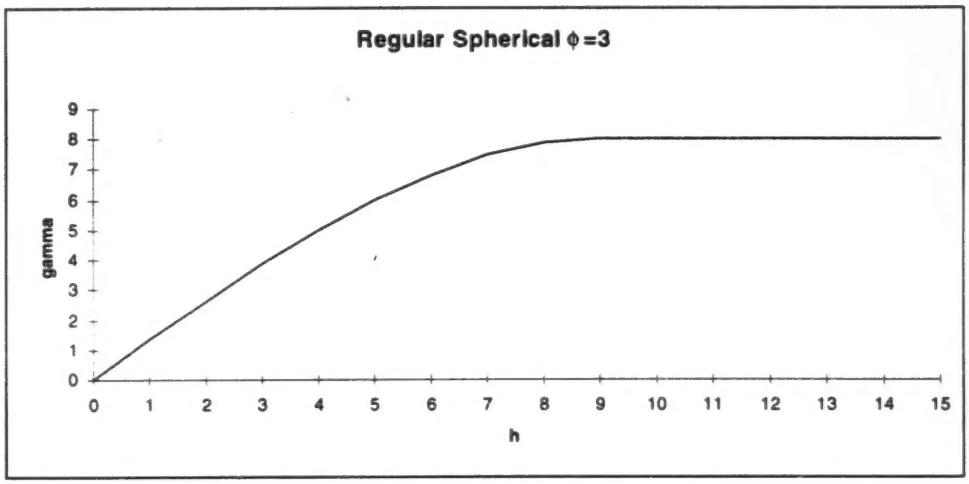
It is important to note that in the graph of the spherical model that $\gamma(h)$ does not approach C asymptotically; it is equal to C when h>a. Like the exponential model , the values are chosen in order to find the best fit. This equation is of this form for the following reasons:

(1) At h=a, $\gamma(h) = C$, and

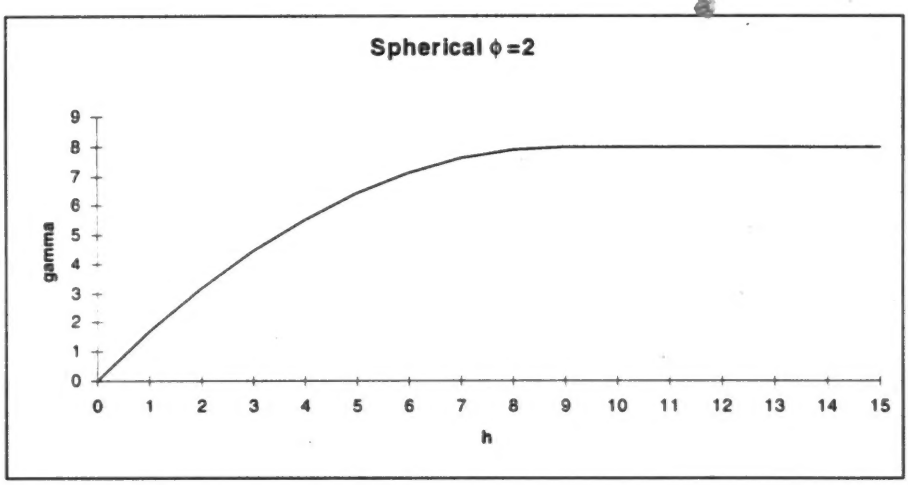
(2) At h=a, the derivative with respect to h of $\gamma(h)$ must be equal to 0.

These conditions must hold in order for the graph of $\gamma(h)$ to approach the point at h=a smoothly. As well, once the value of ϕ has been chosen, once must check for positive definiteness again.

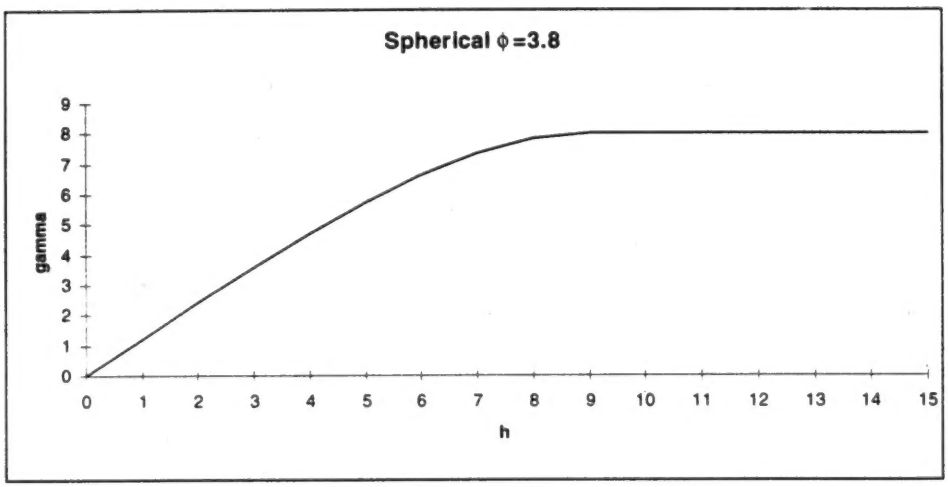
Case(1) Regular Spherical with ϕ = 3:



Case(2) Spherical with ϕ = 2(slightly more hump-shaped):



Case (3) Spherical with ϕ = 3.8(approaches sill less smoothly than regular spherical):

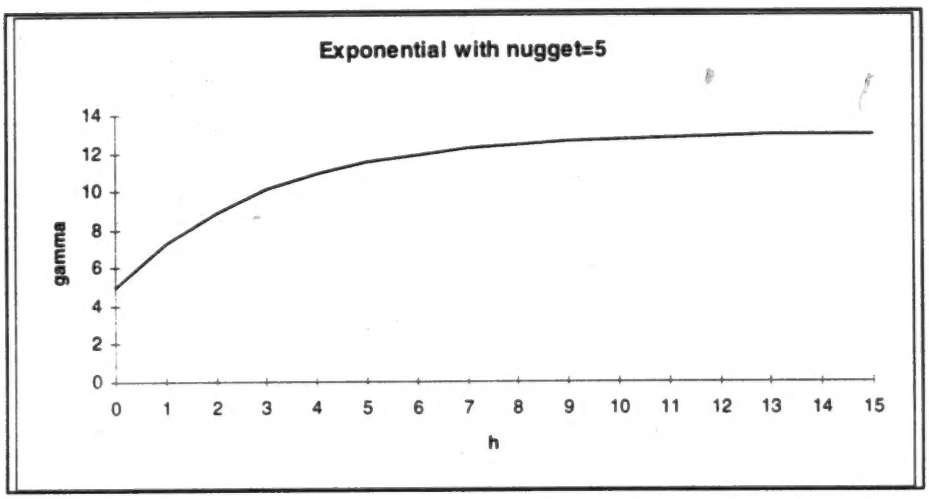


C. NUGGET EFFECT

The nugget effect is defined as the case where $\gamma(h)$ does not approach 0 as h goes to its minimum well-spacing value. It is in many cases the result of spatial variability on a very small scale, and is often no more than a standard semivariogram model with a small range, i.e. $\gamma(h)=C_0[1-\exp(-(h/a))^{\theta}]$, where a is very small relative to the well spacing , and the observed nugget effect is the sill C_0(Hohn, pp. 29-31). It is also chosen arbitrarily in the sense that we inspect the graph of the experimental semivariogram and try and decide on the point where the function crosses the $\gamma(h)$ axis if it is not at 0.

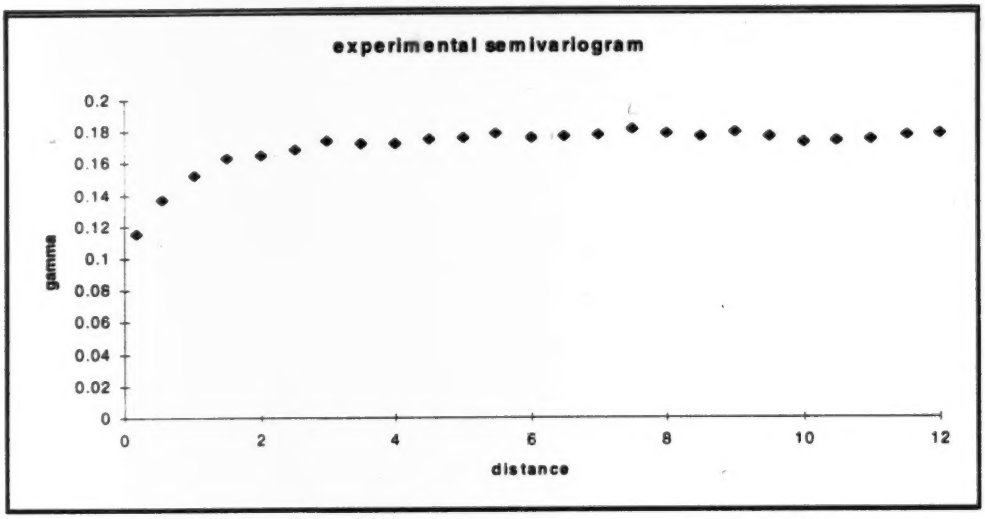
Notice in the plot below how $\gamma(h)$ does not go to 0 as h goes to its minimum well-spacing value.

Case of a Regular Exponential with nugget effect = 5:

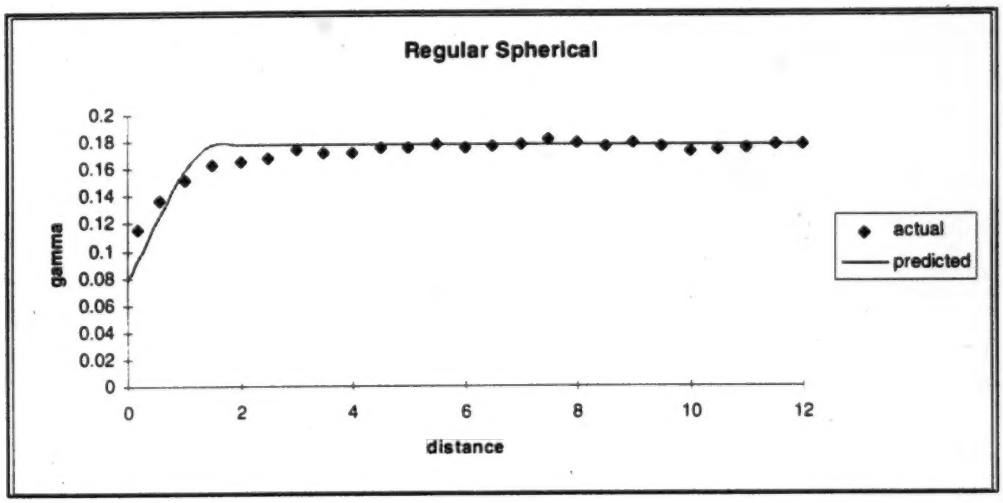


D. EXAMPLE (A)(data on the IP's of reserves in West Virginia in APPENDIX A):

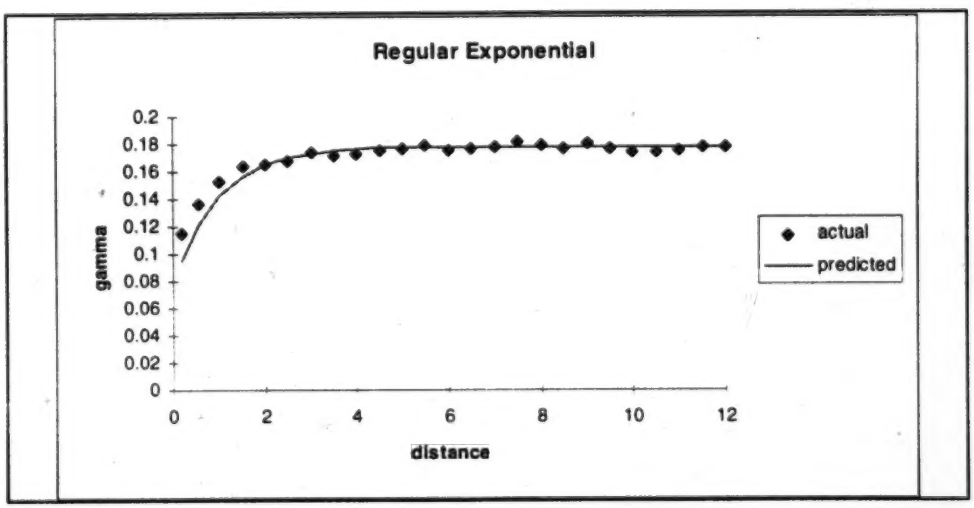
EXPERIMENTAL SEMIVARIOGRAM:



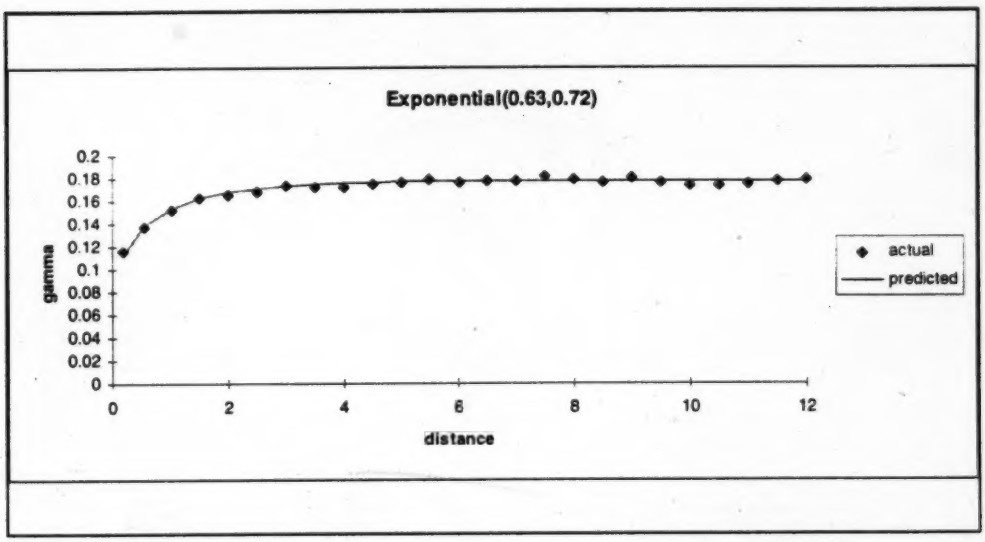
Case(1)Fit a Regular Spherical with a nugget effect of 0.08, a=1.6, C=0.18:



Case(2)Fit A Regular Exponential with C=0.18 and a=0.8:



Case(3)Exponential fit with θ = 0.72, a = 0.63, nugget=0.1, C=0.18:



Notice how this modified exponential model allowed greater latitude in fitting the model to the data and resulted in an excellent fit.

SECTION 5 NESTED MODELS

A. BACKGROUND

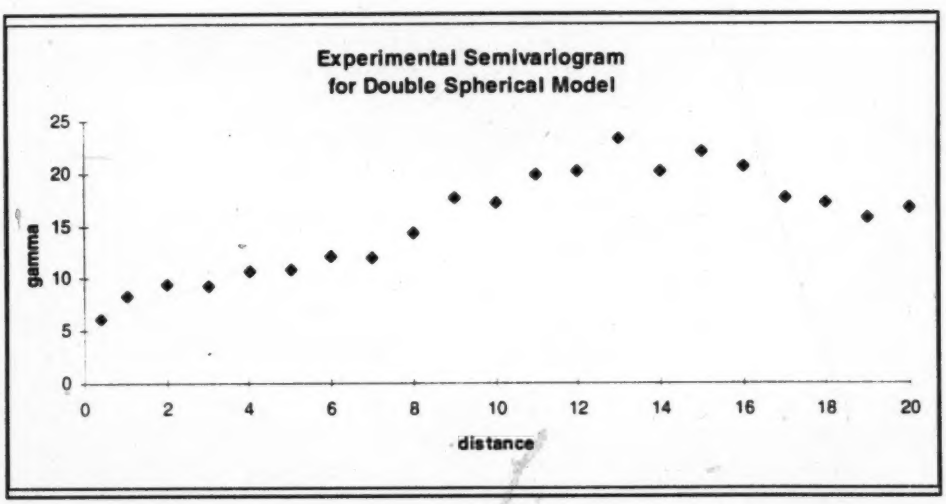
In most real world cases, the final model chosen is not just one of the spherical or the exponential; it is a combination of both(Hohn, pp. 31-33). The general form is the following:

$\gamma(h)=\Sigma_i(\gamma_i(h))$, $i=1,2,.....,I$

if $\gamma_i(h)=C_0$, this model is known as a pure nugget effect.

B. EXAMPLE (B) (data on IP's for reserves in West Virgina in APPENDIX B)

EXPERIMENTAL SEMIVARIOGRAM:

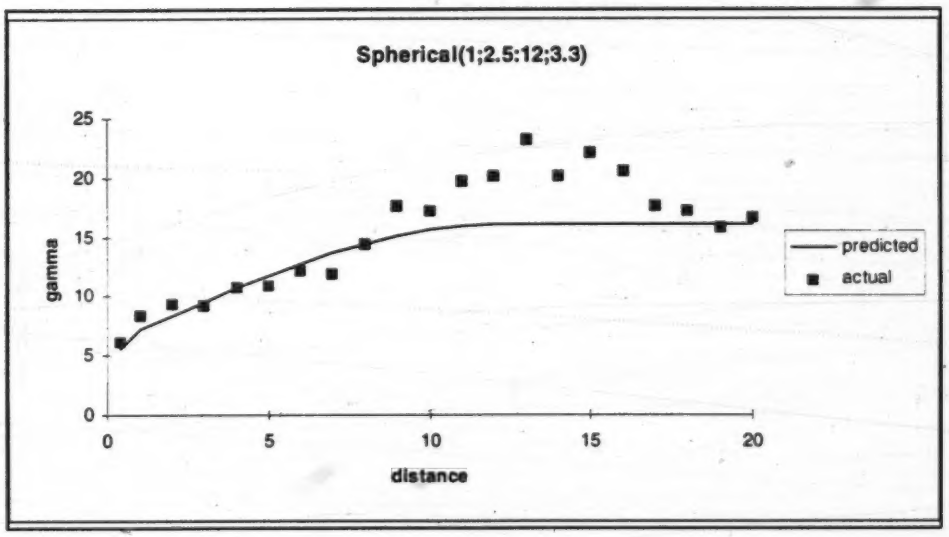


A nested spherical model will be fit to this data set.

We fit a Double Spherical Model to the data, where

(1) Model 1 has a=1 and ϕ=2.5, and

(2) Model 2 has a=12 and ϕ=3.3.



(Note: Due to the lack of a hump-shape of the earlier points, the decision was to try spherical models and not exponential)

SECTION 6 ANISOTROPIES

Anisotropies exist when the term $[z(x)-z(x+h)]^2$ depends on both the value of h and direction; i.e. the orientation of the vector between x and x+h(Journel & Huijbregts, pp. 175-184). The angles are assumed to be measured counter-clockwise from the east. If (x_u, x_v) are the coordinates of a point x in 2 dimensional space, then a corresponding vector h has coordinates (h_u, h_v) and modulus: $|h|=\sqrt{(h_u^2+h_v^2)}$. If anisotropy exists, the semivariogram will be a function of the direction represented by the vector (h_u, h_v).

A. GEOMETRIC ANISOTROPY

A semivariogram $\gamma(h_u, h_v)$ is said to have a geometric anisotropy if there exists a linear transformation such that the following formula holds:

$$\gamma(h_u, h_v) = \gamma'[\sqrt{(h_u'^2 + h_v'^2)}],$$

where $\gamma(h)$ is an anisotropic model and $\gamma'(h)$ is an isotropic model. The variables h_u' and h_v' are defined in the following way:

$h_u' = a_{11}h_u + a_{12}h_v$, where the a_{ij}'s are weighting factors

$h_v' = a_{21}h_u + a_{22}h_v$,

$$[A] = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \underline{h} = (h_u, h_v), \text{ and } \underline{h}' = (h_u', h_v')$$

The matrix form of this equation is:

$$\underline{h}' = [A]\underline{h}, \text{ where } [A] = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

Define Φ to be the angle that the major axis of an ellipse makes with the x-axis, and then the a_{ij}'s satisfy the relations:

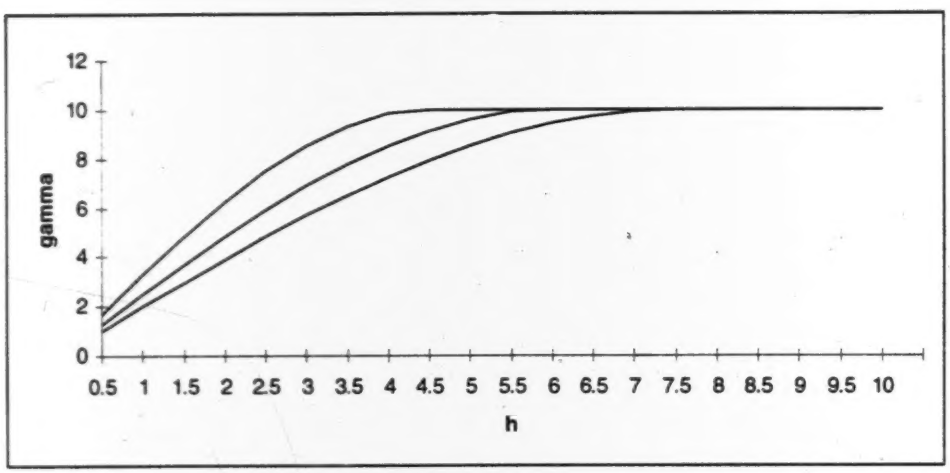
$$a_{11} = \cos^2(\Phi) + k\sin^2(\Phi)$$
$$a_{12} = \sin^2(\Phi) + k\cos^2(\Phi)$$
$$a_{22} = (1-k)\sin(\Phi)\cos(\Phi),$$

where k is the ratio of the anisotropy, i.e. the ratio of the minor axis to the major axis of the ellipse. After finding h_u' and h_v', they are then substituted into the $\gamma(h)$ model. As a result, the direction along the major axis of the anisotropic ellipse is now the x-axis.

If the semivariogram models have a sill, the sill stays constant for all directions if a geometric anisotropy exists. The ranges of the semivariograms vary with direction, but must conform to the following criterion: if the semivariograms are plotted on a graph like the one in EXHIBIT A, the ranges will lie on an ellipse in two dimensions or an ellipsoid in three or more dimensions.

EXHIBIT A

Three semivariograms with ranges 4.5, 6, and 7.5 respectively:



B. ZONAL ANISOTROPY

With a zonal anisotropy each direction is modelled separately in the sense that

$\gamma(h)=\Sigma_i\gamma_i(|h|)$ $i=1,2,.....,n$

Each model is isotropic or anisotropic and is accompanied by the transformation matrix A_i.

General Structural Model for Anisotropies is

$\gamma(h)= \Sigma_j\gamma_j(|h_i|)$

where

$\underline{h}_i= [A_i][\underline{h}]$.

Each $\gamma_j[h]$ has its own anisotropy, and $|C_j(h)|< C_j(0)$, for all j

In the limiting case,

$\gamma[h]=\lim_{N\to\infty}[\Sigma\ \gamma_i(|h_i|)/N]$

i.e. look at it as a relative frequency approach for the infinite different directions.

SECTION 7 HOLE EFFECTS

Sometimes the petroleum geologist encounters the occurrence of significant periodicity. Whereas semivariograms along a structural trend can be fitted by one or two simple models mentioned earlier, a correlation structure may experience highs

and lows as it approaches zero. This process is known as the hole effect. As a result, the correlation between points in the area decreases as the distance h gets larger, but not monotonically as in the case of the two simple semivariogram models(Hohn, pp. 43-44). When fitting these models, it is important to note that since the hole effect varies inversely as the distance h, the models must be chosen in such a way that they dampen cut as h gets larger, since the covariance goes to zero.

Examples of proposed models include the following:

(1) $C(h) = \sin(h)/h$

$\gamma(h) = 1 - \sin(h)/h$

(2) $C(h) = \cos(h)$

$\gamma(h) = 1 - \cos(h)$.

In order to allow more freedom in fitting the models to the data, the following more generalized models have been proposed:

(1) $C(h) = \sin(\omega h)/\omega h$, where ω is an arbitrary natural number

$\gamma(h) = 1 - \sin(\omega h)/\omega h$

(2) $C(h) = \cos(\omega h)$, where ω is an arbitrary natural number

$\gamma(h) = 1 - \cos(\omega h)$.

Case 1 is already damped, but case 2 needs to be damped using

$\gamma(h) = C[1-\{\exp(-(h/a)^6)\cos(\omega h)\}]$,

where C = the sill value.

The amplitude of a hole effect is defined as the ratio of the

minimum value of the covariance divided by the value of the sill.

Amplitude = $|\text{Min}(C(h))|/C(0)$

For model (1), $C(h)=\sin(h)/h$, the amplitude = .212. For a three-dimensional space, the maximum possible amplitude of an isotropic hole effect is 0.212. For the function $C(h)=\sin(h)/h$, $C(h)$ reaches its minimum at $h=3\pi/2$, and $C(3\pi/2) \approx 0.212$. If the observed amplitude is greater than 0.212, this means that the assumed hole effect is illusive. In a one-dimensional model, such as case (2), the maximum amplitude is 1. This is so because cos(h) is positive definite in only one direction. Thus the model in case (2) allows us to fit large hole effects(Journel & Huijbregts, pp 169).

Two more possible models are:

(3) $\gamma(h) = 2 - \sin(\omega h)/\omega h - \cos(\omega h)$

This is dampened using

$\gamma(h)=C[2 - \sin(\omega h)/\omega h - \exp[-(h/a)^{\theta}]\cos(\omega h)]$, ω,θ, arbitrary.

(4) $\gamma(h)= 2-\sin(\omega h)/\omega h - \cos^2(\omega h)$

This is dampened using

$\gamma(h)=C\{2-\exp(-(h/a)^{\theta})\cos^2(\omega h)-\sin(\omega h)/\omega h\}$.

SECTION 8 MODELLING AS A TIME SERIES

A. BACKGROUND

The spatial representation of IP values can also be represented by second order autoregressive processes(Hohn, pp . 52-59). In this case, we simply replace distance h by time t. For time series models, the correlations between data points decrease as t increases. Since the same thing occurs with spatial variability in that the covariance goes to zero as h gets larger, then there is a one to one correspondence between time and distance, allowing us to model the spatial variation in the form of a time series. In order to do so we must first look at how autoregressive(from now on denoted as AR) processes are represented and define all the terms(Cryer, Ch. 4).

z = the variable of interest

z(t) = the value of z at time t

z(t-1)= the value of z at time t-1

z(t-2)= the value of z at time t-2

a(t) = the error term, independent of z(t-1),z(t-2),,,,, etc.

ψ = the autoregressive coefficient

τ = the moving average coefficient

γ_k = the autocovariance for a time series model

$\gamma(h)$ = the semivariogram for a geostatistical model

ρ_k = autocorrelation for a time series model

If second-order stationarity exists, then the covariance and

semivariogram graphs are alternative representations of spatial autocorrelation. The relationship $\gamma(h) = C(0) - C(h)$ implies that there is an inverse relationship between the covariance and the semivariogram. Therefore, while the semivariogram increases with distance h, the covariance decreases. The semivariogram of a regionalized variable displays a sill C(0), which is equal to its sample variance. The existence of this sill implies that a correlogram can be computed using the following formula:

$\rho(h) = C(h)/C(0) = 1 - \gamma(h)/C(0)$

If this formula is rewritten, we get $\gamma(h) = C(0)(1-\rho(h))$, which gives the semivariogram corresponding to a given autoregressive process. At this point it should be pointed out that the ρ_k for a time series process is equivalent to the $\rho(h)$ for a semivariogram process in the sense that they are both measures of correlation. The difference being that for the time series model the correlation is measured over time and for the semivariogram model the correlation is measured over distance. Therefore if ρ_k for a time series model can be found, then its semivariogram equivalent can be determined, and used to model a particular semivariogram.

An AR(2) process has the following representation:

$z(t) = \psi_1 z(t-1) + \psi_2 z(t-2) + a(t)$ {1}

The characteristic polynomial of z is:

$\psi(x) = 1 - \psi_1 x - \psi_2 x^2$, and ψ always has 2 roots(possibly complex)

A stationary solution to {1} exists if and only if the roots

exceed unity in absolute value. For the AR(2) case,

$x = [\psi_1 \pm \sqrt{(\psi_1^2 - 4(1)(-\psi_2))}]/(-2\psi_2)$, using the quadratic formula.

These roots are greater than 1 in absolute value if and only if

(a) $\psi_1 + \psi_2 < 1$,

(b) $\psi_2 - \psi_1 < 1$, and

(c) $|\psi_2| < 1$, all hold simultaneously.

(1) $\gamma_k = \psi_1\gamma_{k-1} + \psi_2\gamma_{k-2}$ = autocorrelation function

(2) $\rho_k = \psi_1\rho_{k-1} + \psi_1\rho_{k-2}$,

and (2)=(1)/γ_0, therefore $\rho_1 = \psi_1/(1-\psi_2)$, which implies that

$\rho_2 = \psi_1\rho_1 + \psi_2 = [\psi_2(1-\psi_2) + \psi_1^2]/(1-\psi_2)$, and we can then obtain successive values of ρ_k from (2).

If the roots of the characteristic polynomial are complex then,

$\rho_k = \sqrt{[(-\psi_2)\sin(Tk + \Psi)/\sin(\Psi)]}$

$\cos[T] = \psi/(2\sqrt{(-\psi_2)})$,

$\tan[\Psi] = \tan[T](1-\psi_2)/(1+\psi_2)$.

If the roots are equal,

$\rho_k = [1 + k(1+\psi_2)/(1-\psi_2)](\psi_1/2)^k$, k=0,1,2,,,,,,.

An exponential semivariogram results from an AR(1) of form $z(x_i) = \psi z(x_{i-1}) + \epsilon_i$, with ψ=1. If ψ<1, the resulting semivariograms are similar to the general exponential model, $\gamma(h) = C(0)[1-\exp(-(h/a)^\theta)]$. If ψ is near 0.9, then there will be a large range. Smaller values of ψ will reduce the range. The equivalence of the ARMA(1,0) model to the exponential model

for γ will now be demonstrated. The autocorrelations are

$\rho(0) = \psi^0 = 1$

$\rho(1) = \psi\rho(0) = \psi^1$

$\rho(2) = \psi\rho(1) = \psi^2$

$\rho(3) = \psi\rho(2) = \psi^3$

.

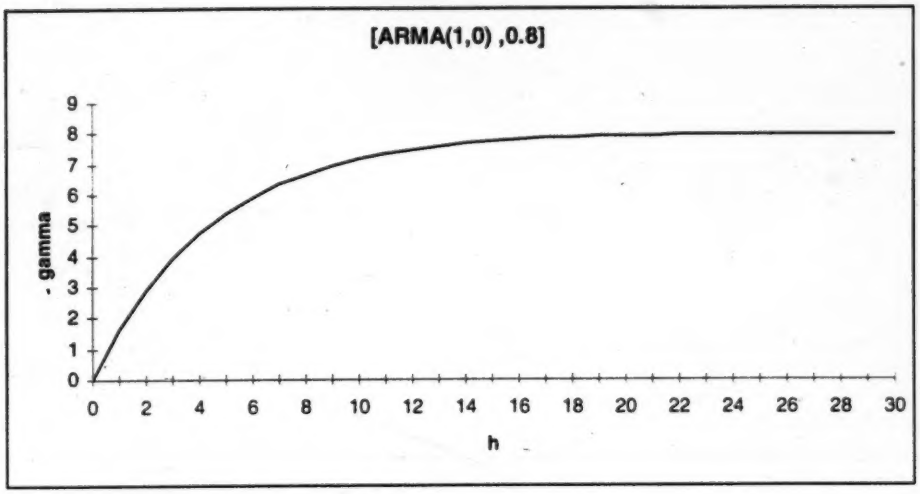
.

$\rho(h) = \psi^h$.

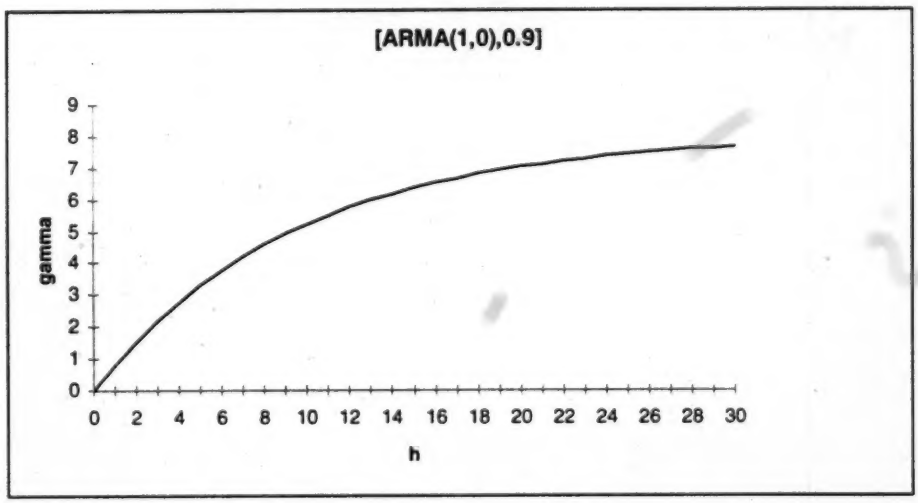
Therefore $\gamma^*(h) = C(0)[1-\psi^{h'}]$, this is the exponential model with $\psi^{h'} = \exp(-(h/a)^\theta)$, where $h' = h^\theta$, and therefore $a = -1/[\ln(\psi)]^{1/\theta}$.

ARMA(2,0) models give greater latitude in model building. If $\psi_1^2 + 4\psi_2 > 0$, the resulting graph is a Gaussian type curve. If $\psi_1^2 + 4\psi_2 < 0$, the resulting graph has a hole effect due to the pseudoperiodicity in the data(Cryer, pp 68). At this point it is important to note that in order to preserve stationarity, that $\psi_1 + \psi_2 < 1$, must hold.

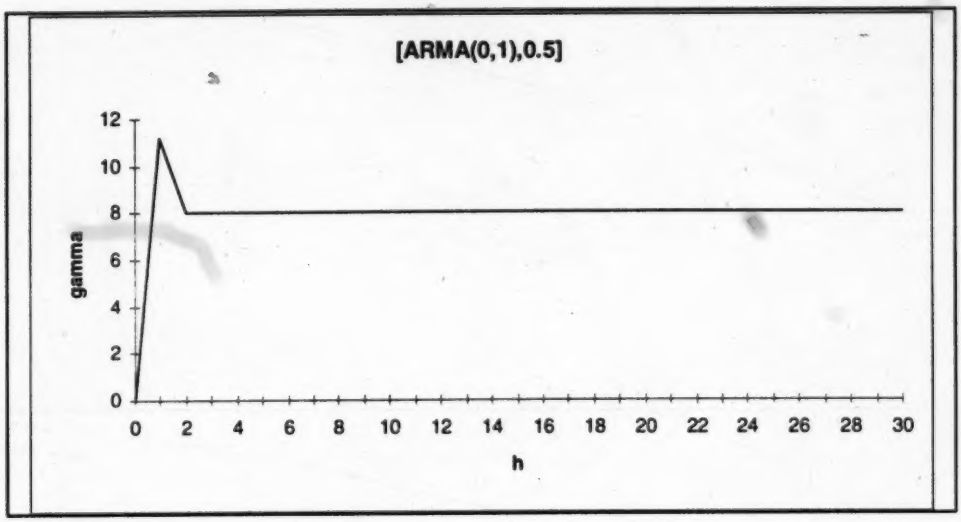
B. CASE(1):ARMA(1,0) with $\psi = 0.8$



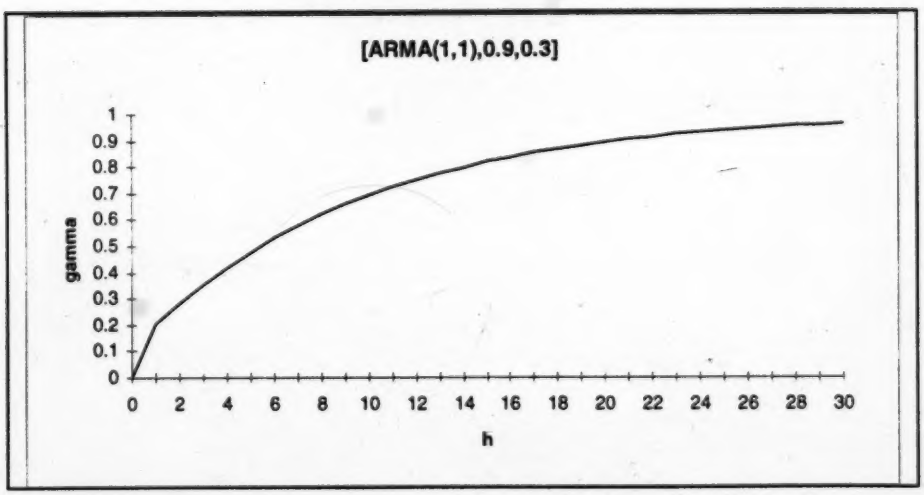
C. CASE(2): ARMA(1,0) with $\psi = 0.9$



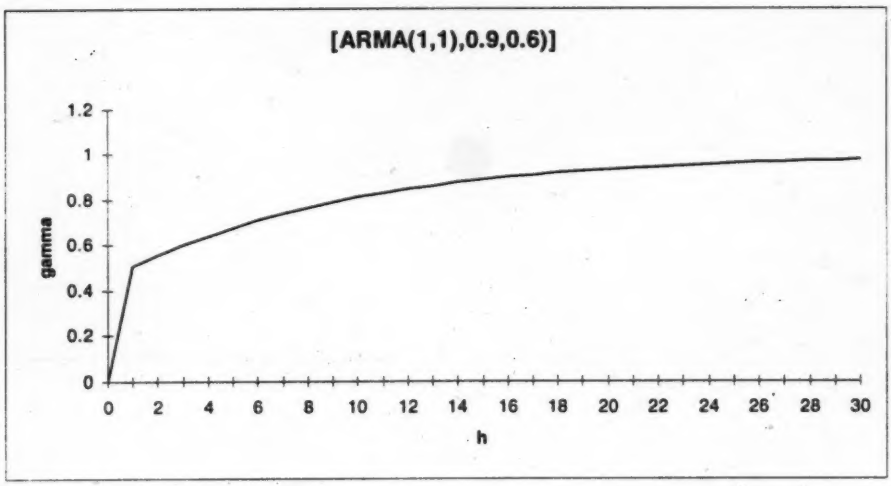
D. Case(3) ARMA(0,1) with τ=0.5



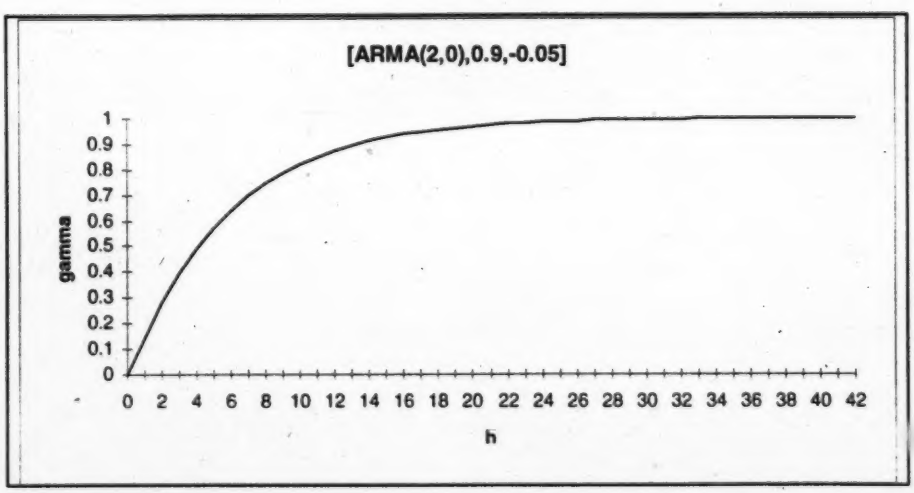
E. Case(4) ARMA(1,1) with ψ=0.9 and τ=0.3



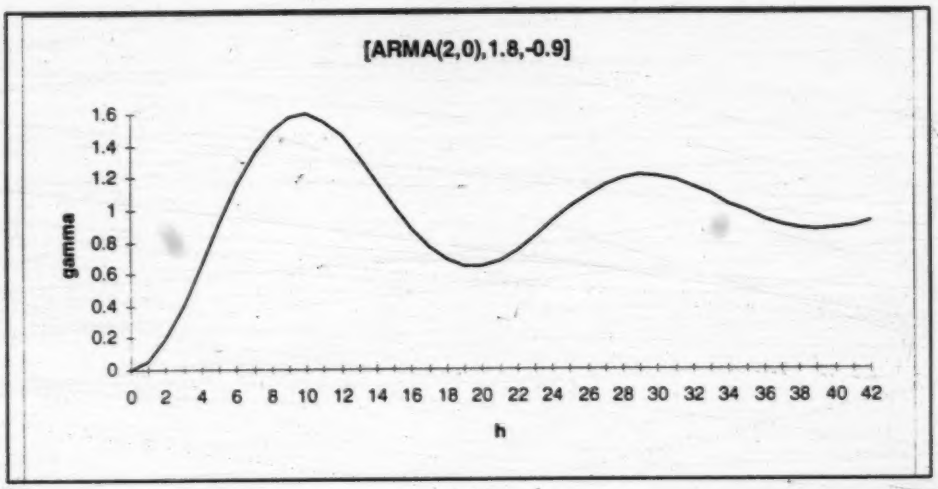
F. Case(5) ARMA(1,1) with ψ=0.9 and τ=0.6



G. Case(6) ARMA(2,0) with ψ_1 = 0.9 and ψ_2 = -0.05

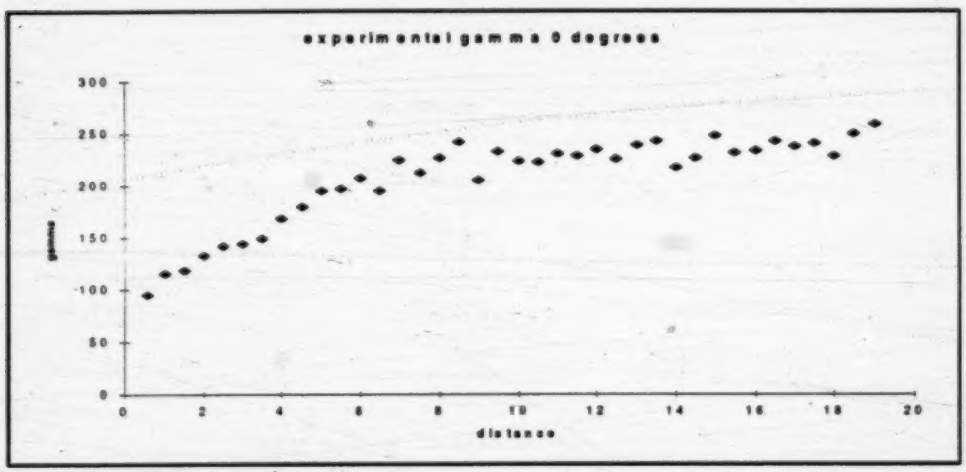


H. CASE(7)ARMA(2,0) with $\psi_1 = 1.8$ and $\psi_2 = -0.9$



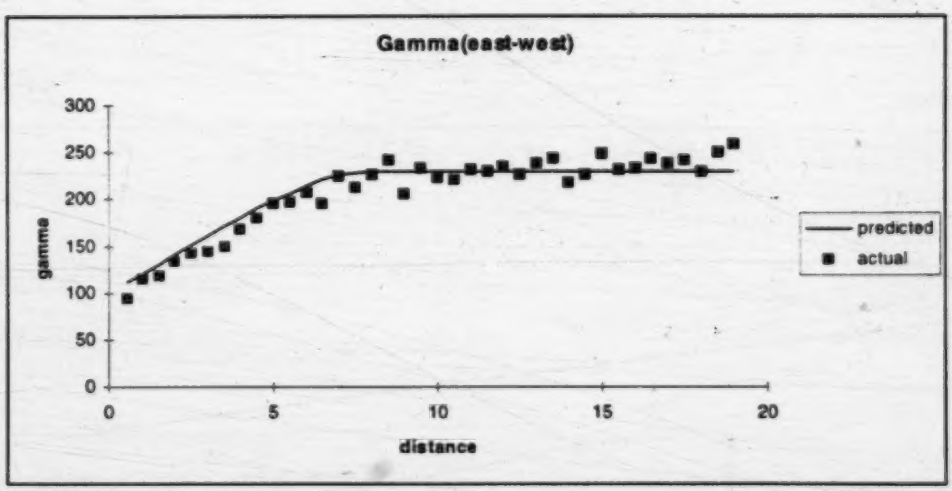
SECTION 9 EXAMPLE C: EXAMPLE OF ANISOTROPIES IN THE 4 BASIC DIRECTIONS(data on IP's from reserves in West Virginia in APPENDIX C)

EXPERIMENTAL SEMIVARIOGRAM FOR EAST-WEST DIRECTION:

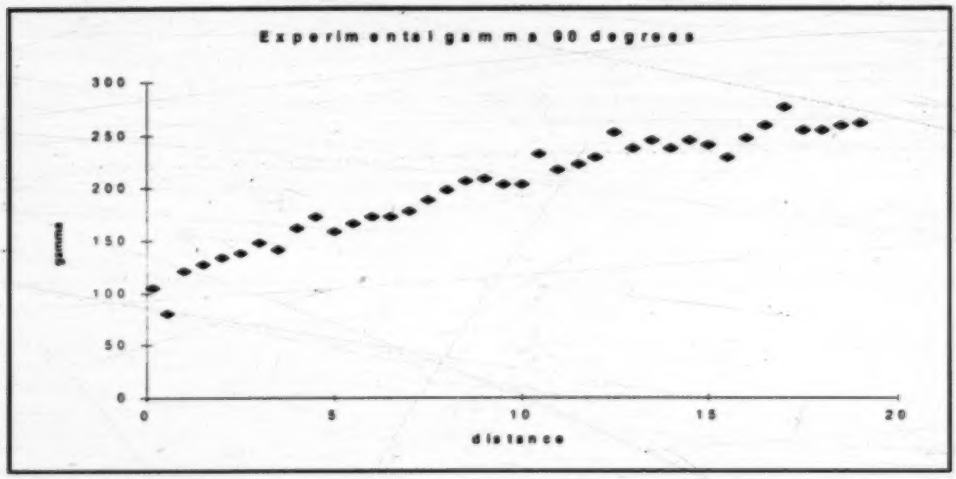


Fitted Semivariogram for east-west direction:

Nested model: pure nugget effect plus an spherical with Φ =8.

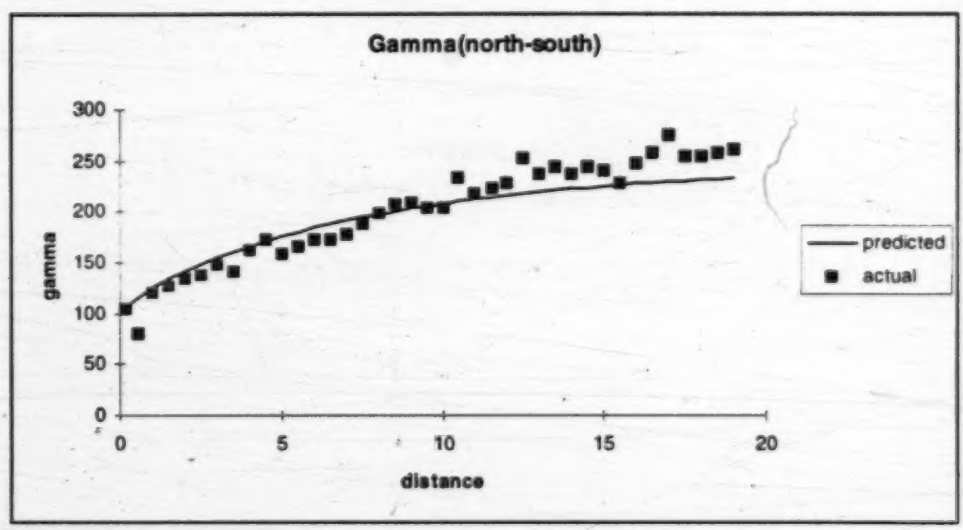


EXPERIMENTAL SEMIVARIOGRAM FOR NORTH-SOUTH DIRECTION:

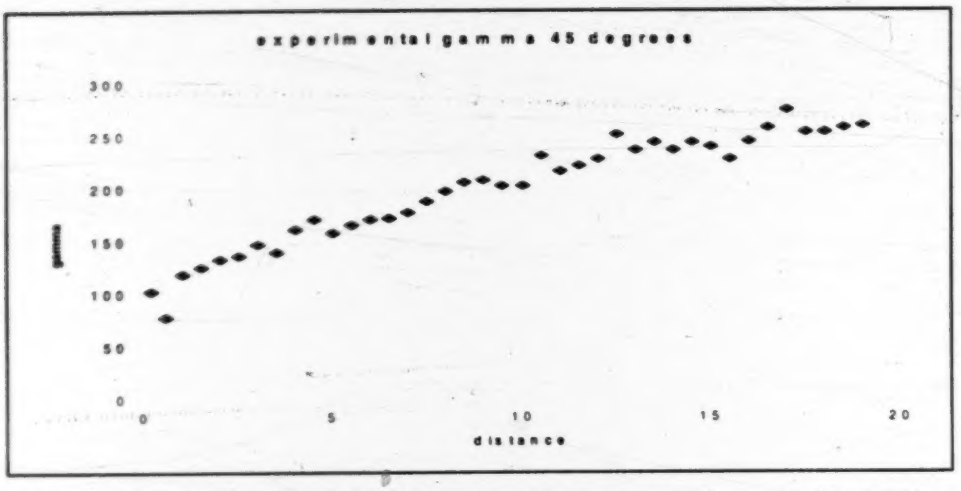


Fitted Semivariogram for the north-south direction:

Nested Model: Pure nugget effect plus Exponential with θ=0.85.

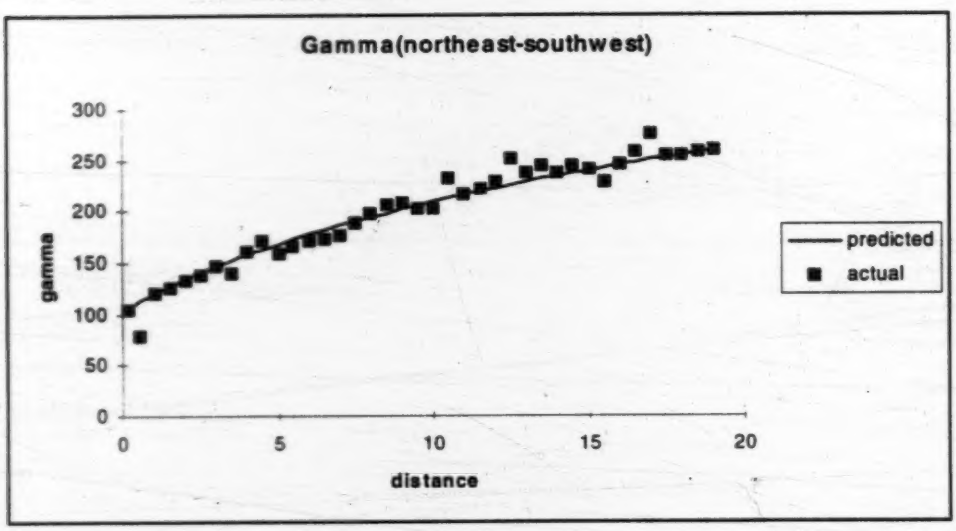


EXPERIMENTAL SEMIVARIOGRAM FOR NORTHEAST-SOUTHWEST DIRECTION:

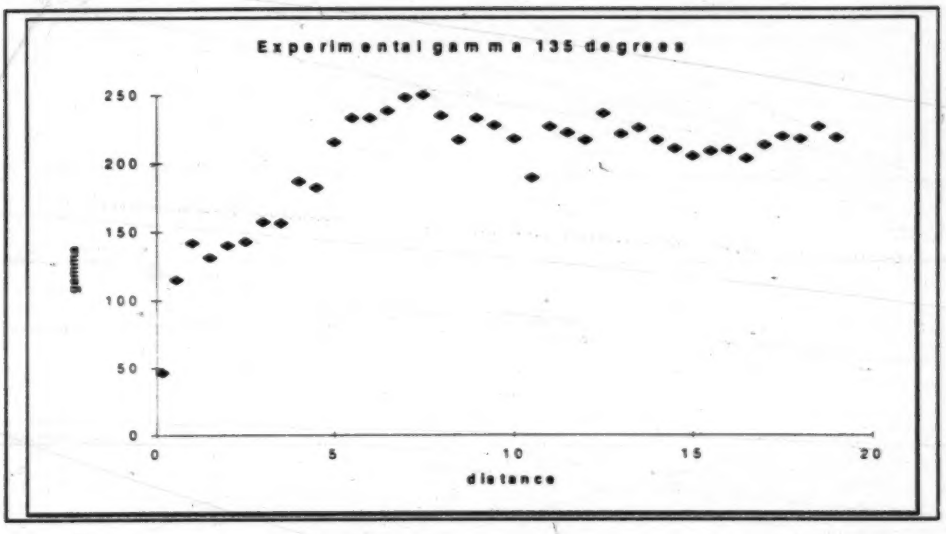


Fitted Semivariogram for the northeast-southwest direction:

Nested Model: Pure nugget effect plus an Exponential with $\theta=0.8$.

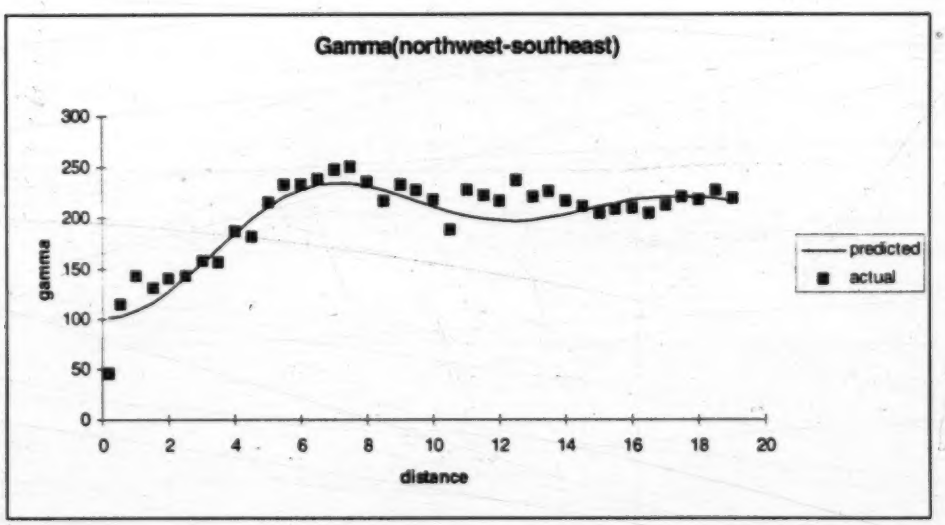


EXPERIMENTAL SEMIVARIOGRAM FOR NORTHWEST-SOUTHEAST DIRECTION:



Fitted semivariogram for the northwest-southeast direction:

Nested Model: Pure nugget effect plus a spherical with $\phi=4$ and a hole effect model of the form $\gamma(h)=1-(\sin(0.628h))/(0.628h)$.



SECTION 10 PROPORTIONAL EFFECT AND QUASI-STATIONARITY

Let $x = (x_u, x_v)$ and $x' = (x_u', x_v')$ be two points on a particular grid(Journel & Huijbregts, pp.186). In the absence of the hypothesis of stationarity,

(1) The expectation of $z(x)$ depends on the location of the regionalized variable on the grid,

i.e. it depends on the coordinates of x

$$E\{z(x)\} = \mu(x_u, x_v).$$

(2) Both the semivariogram $\gamma(x,x')$ and covariance $C(x,x')$ depend on the locations of x and x', i.e. $(1/2)E\{[z(x) - z(x')]^2\} = \gamma(x,x')$.

Under the hypothesis of quasi-stationarity we have

(1) The expectation of $z(x)$ is quasi-constant over a limited number of neighbourhoods and $\mu(x) \approx \mu(x') \approx \mu(x_0)$ when x and x' are inside the neighbourhood $V(x_0)$, which is centred at a point x_0.

(2) Within $V(x_0)$, γ and C depend only on $h = x - x'$ and not on the locations of x and x'; but they are dependent on neighbourhood $V(x_0)$ in which they are found, i.e. on the location of x_0.

Thus $\gamma(x,x') = \gamma(x-x', x_0) = \gamma(h, x_0)$, $\forall\ x, x' \in V(x_0)$.

A. PROPORTIONAL EFFECT

Let $V(x_0)$ and $V(x_0')$ be two neighbourhoods centred on x_0 and x_0'. Assume they are quasi-stationary on x_0 and x_0' respectively and that $\gamma(h,x_0)$ and $\gamma(h,x_0')$ are the semivariograms defined on these 2 neighborhoods(Journel & Huijbregts, pp.187-190). The two γ's can be made to coincide after being multiplied by a factor which is a function of the experimental means $\mu^*(x_0)$ and $\mu^*(x_0')$, therefore we assume the existence of a stationary model $\gamma_0(h)$ that is independent of the neighbourhood $V(x_0)$ and such that

$$\gamma(h,x_0)=f[\mu^*(x_0)]\gamma_0(h).$$

B. CASES OF A PROPORTIONAL EFFECT(Journel & Huijbregts pp 188-189):

(1) Direct Case:

$\gamma(h,x_0)/f[\mu^*(x_0)]= \gamma(h,x_0')/f[\mu^*(x_0')]$, where

$\gamma(h,x_0)$ and $\gamma(h,x_0')$ differ by a proportional effect in a direct manner. This would be the case if the experimental semivariogram varies directly as the corresponding experimental mean. This would be true if z(x) has a lognormal-type histogram.

(2)Inverse Case:

Here $\gamma(h,x_0)f[\mu^*(x_0)]= \gamma(h,x_0')f[\mu^*(x_0')]$.

$\gamma(h,x_0)$ and $\gamma(h,x_0')$ differ by a proportional effect in an

inverse manner. This would be the case if the sills of the experimental semivariograms vary inversely as the corresponding experimental mean. This would be true if z(x) had an inverse lognormal-type histogram.

(3) Case where a proportional effect in $(\mu^*)^r$ exists, r is arbitrary:

1. Direct: $f[\mu^*(x_0)]=[\mu^*(x_0)]^r$
2. Inverse: $f[\mu^*(x_0)]=[\ A - \mu^*(x_0)]^r$,

where A = the maximum value that μ^* can assume.

Here is an example of a case where two semivariograms have nugget effects.

$\gamma(h,x_0)=[f(\mu^*(x_0))+N_0]\gamma_0(h)$, and $\gamma(h,x_0')=\gamma_0(h)[f[\mu^*(x_0')]+N_0']$, where N_0, N_0' are nugget effects.

Therefore in the case with $N_0=0$,

$\gamma(h,x_0)/f[\mu^*(x_0)]=\gamma(h,x_0')/f[\mu^*(x_0')]$

and in the general case

$\gamma(h,x_0)/[f(\mu^*(x_0)+N_0]=\gamma(h,x_0')/[f(\mu^*(x_0'))+N_0']$

SECTION 11 AUXILIARY FUNCTIONS

Sometimes we may want to find the average value of a semivariogram in a one-dimensional deposit over a rectangular block of length L, instead of finding every value of $\gamma(h)$ at

each point in this area(Rendu, pp. 30-43). This is done by means of what are known as auxiliary functions. These are functions related to the actual semivariograms by means of certain mathematical formulas, which are used separately to derive these functions for the exponential and spherical models. The rectangular block is defined as having corners A,B,C,D,E,F,G,H, and two square faces ACEG and BDFH. As mentioned earlier, the length of the rectangular block P is L, i.e the distance from the side ACGE to BDHF, and the height and width are both of length d.

The two types of auxiliary functions that will be examined are

(1) $\chi(L;d^2) = \bar{\gamma}(ACEG;P)$, the mean value of $\gamma(h)$ between one square face and P, and

(2) $F(L;d^2) = \bar{\gamma}(P;P)$, the mean value of $\gamma(h)$ within P.

A. <u>MATHEMATICAL REPRESENTATION OF THE AUXILIARY FUNCTIONS</u>

$$\chi(L) = (1/L) \int_0^L \gamma(h)dh$$

$$F(L) = (2/L^2) \int_0^L \int_y^L \gamma(h-y)dhdy$$

B. GENERALIZED EXPONENTIAL MODEL

The auxiliary functions F(L) and $\chi(L)$ for the generalized exponential model will now be computed.

$$F(L) = (2/L^2) \int_0^L \int_y^L (1-\exp(-((h-y)/a)^{\theta}))\,dh\,dy$$

$$=(2/L^2) \int_0^L \int_0^h (1-\exp(-((h-y)/a)^{\theta})\,dy\,dh$$

let $u=((h-y)/a)\theta$,

therefore $du = (\theta/a)u^{(\theta-1)/\theta}$.

Let $\alpha = (1/\theta)$,

therefore $du = (1/\alpha a)u^{(1-\alpha)}$, and

$$F(L)= -a^2 \int_0^L \int_{au^{\alpha}}^{L} u^{\alpha-1}\exp(-u)\,du\,dh$$

$$= -a^2 \int_0^L (L-au^{\alpha})u^{\alpha-1}\exp(-u)\,du$$

$$= -a^2L\Gamma(\alpha)+ a^3 \int_0^L u^{2\alpha-1}\exp(-u)\,du$$

$= -a^2[\Gamma(\alpha)-a\Gamma(2\alpha)]$,

$= -a^2[\Gamma(1/\theta)-a\Gamma(2/\theta)]$

where Γ is the standard gamma function.

Now $\chi(L) = (1/L)\int_0^L 1-\exp(-(L/a)^{\theta})dh$,

Let $u = (h/a)^{\theta}$,

therefore $du = (\theta/a)u^{(\theta-1)/\theta}$.

Let $\alpha = 1/\theta$,

therefore $du = (\theta/a)u^{1-\alpha}dh$, and

$dh = (a/\theta)u^{\alpha-1}du$, so

$$\chi(L) = (\alpha a/L)\int_0^g u^{\alpha-1}du - \int_0^g u^{\alpha-1}\exp(-u)du,$$

$$= 1- (\alpha a/L)\int_0^g u^{\alpha-1}\exp(-u)du,$$

where $g = (L/a)^{(1/\alpha)}$, now

$\chi(L) = 1 - (\alpha a/L)\gamma(\alpha,(L/a)^{(1/\alpha)})$,

$= 1 - (a/\theta L)\gamma(1/\theta,(L/a)^{\theta})$,

where $\gamma(\alpha,x)$ is the incomplete gamma function(Magnus, Oberhettinger, Tricomi, pp 387).

C. GENERALIZED SPHERICAL MODEL:

The auxiliary functions $\chi(L)$ and $F(L)$ for the generalized spherical model will now be computed.

$$\gamma(h) = (\phi/(\phi-1))[h/a] - (1/(\phi-1))[h/a]^{\phi} \quad , h \le a$$

$$= 1 \quad , h > a$$

$$\chi(L) = (1/L) \int_0^L \{(\phi/(\phi-1))[h/a] - [1/(\phi-1)][h/a]^{\phi}\}dh$$

$$= (1/L)[(\phi/(2\phi-2))(h^2/a) - [1/(\phi-1)(\phi+1)][h^{\phi+1}/a^{\phi}]]$$

$$= [\phi/(2\phi-2)](L/a) - [1/(\phi-1)(\phi+1)][L/a]^{\phi}, \; L<a,$$

and if $L \ge a$

$$\chi(L) = [1/L][\int_0^a (\phi/(\phi-1))[h/a] - (1/(\phi-1))[h/a]^{\phi}dh + \int_a^L dh$$

$$= [1/L][(\phi/(2\phi-2))(h^2/a) - (1/(\phi-1)(\phi+1))(h^{\phi+1}/a^{\phi})]_0^a + (1/L)[L-a]$$

$$= (1/L)[(\phi a - \phi^2 a)/(\phi-1)(\phi^2-1) - L]$$

$$= 1 - \phi(\phi^2-1)[a/L], \; L \ge a.$$

$$F(L) = [2/L^2] \int_0^L \int_y^L \{(\phi/(\phi-1))[(h-y)/a] - (1/(\phi-1))[(h-y)/a]dhdy$$

$$= [2/L^2] \int_0^L [\phi(L-y)^2/2(\phi-1)a] - (L-y)^{\phi+1}/(\phi-1)(\phi+1)(\phi+2)a^{\phi}dy$$

$$= [2/L^2]\{(\phi L^3/6(\phi-1)a) - L^{\phi+2}/(\phi-1)(\phi+1)(\phi+2)a^{\phi}\}$$

$$= \phi L/(3(\phi-1)a) - (2/(\phi-1)(\phi+1)(\phi+2))(L/a)^{\phi}, \text{ if } L<a,$$

and if $L \geq a$, then

$$F(L) = [2/L^2][\int_0^a \int_y^a (\phi/(\phi-1))[(h-y/a] - (1/(\phi-1))[(h-y)/a]^2 dhdy$$

$$+ \int_a^L \int_y^L dhdy$$

$$=(2/L^2)[y^2/2 - aL + a^2/2] + (2/L^2) \int_0^a \{\phi(a-y)^2/(2a(\phi-1)) -$$

$$(a-y)^{\phi+2}/(\phi-1)(\phi+1)a^{\phi}\}dy$$

$$=[2/L^2][y^2/2 - aL + a^2/2 - \phi a/6(\phi-1) - a^2/(\phi-1)(\phi+1)(\phi+2)]$$

$$=1 + [(\phi^2-1)(\phi+2)a^2 - 2a^2]/L^2(\phi-1)(\phi+1)(\phi+2) - 2a/L - \phi a/L^2 3(\phi-1)$$

SECTION 12 KRIGING

A. BACKGROUND

Named after the South African geologist D.G. Krige, kriging is a procedure that estimates the values of the geological variable under study at unknown sites on a chosen area V. It is

a local estimation procedure which provides the best linear unbiased estimator of the variable being studied(Journel & Huijbregts, pp. 304). It is limited to the linear class of estimators, since only the second moment (covariance or semivariogram)of the regionalized variable is required. The parameters are then chosen in order to ensure that the estimates are unbiased. With regard to oil reserves, a predetermined grid is set up for the area under study. Then in order to model the variation of the IP's across the grid, a semivariogram is chosen in the manner explained earlier. Kriging is then performed by using the covariances of the known IP's in order to estimate the IP's of potential well sites on the grid.

Let $\gamma(h)$ be the semivariogram chosen.

Let $z(x)$ be the IP at a point x.

The conditions for second-order stationary of $z(x)$ are

(1) $E\{z(x)\}= \mu$, an unknown constant,

(2) $C(h)= E\{z(x+h)z(x)\} - \mu^2$, the centred covariance,

(3) $\gamma(h)= E\{[z(x+h)-z(x)]^2\}/2$, the semivariogram.

Either of these second order moments is assumed to be known.

Let z_k^* = the kriged estimator of the average IP over some area V. It is a linear combination of n values of the regionalized variable, where n is the number of IP's that are used in the kriging process. The kriging estimate is $z_k^* = \Sigma_i \lambda_i z_i$, $i=1,2,,,,,,n$. The z_i's are the mean IP's $z_{vi}(x_i)$ defined on the supports v_i, which are centred on each point x_i. These n supports can differ from each other. Under the hypothesis of

stationarity, the expectation of each z_i is $E\{z_i\}= \mu$, for all i. The λ_i's are the weights calculated in order to ensure that our three conditions above are satisfied. As well, if $\Sigma_i\lambda_i=1$, then this will ensure that the estimate is unbiased and that $E[z_k^*] = \mu\Sigma_i\lambda_i = \mu = E(z_v)$, where z_v is the unknown average IP for the area V, and that $E[z_v-z_k^*]=0$.

The second condition that states that the estimation variance has to be minimized. This means that we must minimize

$$E[(z_v-z_k^*)^2]=E[z_v^2]-2E[z_vz_k^*] + E[(z_k^*)^2] = \bar{C}(V,V) - 2\Sigma_i\lambda_i \bar{C}(v_i,V)$$

$$+ \Sigma_i\Sigma_j\lambda_i\lambda_j \bar{C}(v_i,v_j),\ i=1,..n\ ,\ j=1,...,m,$$

where $\bar{C}(A,B)$= the average covariance between each point in an area A and each point in an area B.

B. KRIGING SYSTEM

There are n+1 linear equations with n+1 unknowns obtained by equating each of their partial derivatives of $E\{[z_v-z_k^*]^2\} - 2\delta\Sigma_i\lambda_i$ to zero(Journel & Huijbregts pp.304-308),

$$\frac{\partial}{\partial\lambda_i}\{E[(z_v-z_k^*)^2]] - 2\delta\Sigma_j\lambda_j\}=0,\quad i=1,2,.....,n,$$

where δ is a Lagrange multiplier.

The n weights are to be calculated from this.

The result is the kriging system:

(1) $\Sigma_i \lambda_i \bar{C}(v_i, v_j) - \delta = \bar{C}(v_j, V)$ for all $j = 1$ to n

(2) $\Sigma_i \lambda_i = 1$

And the minimum estimation variance is given by

$$\sigma_k^2 = \bar{C}(V,V) + \delta - \Sigma_i \lambda_i \bar{C}(v_i, V).$$

C. MATRIX FORM

$$W = \begin{bmatrix} \bar{C}(v_1, v_1) & & \bar{C}(v_1, v_n) & 1 \\ \bar{C}(v_2, v_1) & & \bar{C}(v_2, v_n) & 1 \\ & \cdot & & \\ & \cdot & & \\ & \cdot & & \\ \bar{C}(v_n, v_1) & & \bar{C}(v_n, v_n) & 1 \\ 1 & 1 & & 0 \end{bmatrix}$$

$$[B] = \begin{bmatrix} \bar{C}(v_1, V) \\ \bar{C}(v_2, V) \\ . \\ . \\ . \\ \bar{C}(v_n, V) \\ 1 \end{bmatrix} \qquad \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ . \\ . \\ . \\ \lambda_n \\ \delta \end{bmatrix}$$

Then $[W]*[\lambda] = [B]$ and therefore $[\lambda]=[W]^{-1}[B]$.

Explanation of the terms.

[W] is the covariance matrix for the points $v_1, v_2, \ldots\ldots, v_n$. The distance between well sites i and j is calculated giving h_{ij} as well as the direction of the vector described by two samples if

an anisotropic situation exists. The matrix [B] is composed of the elements $\bar{C}(v_i,V)$, where $\bar{C}(v_i,V)$ is the mean value of the covariance function C(h) when the domains V and v_i are independently described by the two extreme points of the vector h. Due the one to one correspondence between the covariance and the semivariogram, the value of $C(v_i,v_j)$ is obtained from $\gamma(v_i,v_j)$, which is determined from the chosen semivariogram model. Although the calculation of point-to-block covariance should be done using integration, in practice it is done using discrete summation. The covariance between each point in a block and the sample v_i is calculated, and then an average value is determined and used for $\bar{C}(v_i,V)$.

If, for example, V is approximated with k points, then k values of $\bar{C}(v_i,v_j)$ are calculated.

D. KRIGED ESTIMATES

The weights λ depend on 4 factors:

(1) The size and shape of the block to be estimated. Since the average covariance of points in the block varies inversely as the block size, the estimation variance will reach its maximum

when estimating a variable at a point. For this reason, $\bar{C}(V,V)$ can be used as a measure of the size and shape of the block.

(2) $\bar{C}(V,v_j)$ = covariance between points in the block and the control points. But since distance varies directly as estimation variance and inversely as the average covariance, it is essentially a measure of the distance between the points in the block and the control points.

(3) $\bar{C}(v_i,v_j)$ = the covariance between the control wells, but for the same reasons as in (2), it is also a measure of the distance and configuration between control wells.

(4) The quality and value of the estimate depend on the γ model. The kriging system of equations and kriging variance depend on the size and shape of V, the distance between V and the control points and the configuration of the control points and γ.

E. INFLUENCE OF A NUGGET EFFECT (Journel & Huijbregts, pp. 310-312)

$C(h) = N_0 + C_1(h)$, where N_0 is the nugget effect, and $C_1(h)$ is the usual covariance at distance h.

$$\bar{C}(v_\alpha, v_\beta) = \bar{C}_1(v_\alpha, v_\beta) + A/v_\alpha, \quad v_\beta \subset v_\alpha$$

$$= \bar{C}_1(v_\alpha, v_\beta), \quad v_\alpha \cap v_\beta = \emptyset,$$

where A is a constant given by $A = \int N_0 dh$.

The kriging system is:

$$\lambda_\alpha[A/v_\alpha + \bar{C}_1(v_\alpha, v_\beta)] + \Sigma_{\alpha,\beta}[\lambda_\alpha \bar{C}_1(v_\alpha, v_\beta)] - \delta = \bar{C}_1(v_\alpha, V), \text{ for all } \alpha.$$

$\Sigma_\beta \lambda_\beta = 1$, if all supports V, v_α, v_β, are disjoint. The minimum estimation variance provided by kriging is then given by

$$\sigma_k^2 = A/V + \bar{C}_1(V,V) + \delta - \Sigma_\alpha \lambda_\alpha \bar{C}_1(v_\alpha, V)]$$

If all the supports are not disjoint, then each of the $\bar{C}$ terms in the kriging system must be calculated separately.

SECTION 13 MULTIVARIATE CASE

A. COREGIONALIZATION

Coregionalization* occurs when there are m regionalized variables $z_1,\ldots,z_m$ and n sample locations(Hohn, pp.140-141). This leads to the multivariate extension of kriging known as cokriging. In cokriging, the analysis performed is essentially the same as in kriging in that the IP values of potential undrilled well sites on the grid are predicted using the semivariogram models that have been chosen. Only now the values of other variables are included, such as natural gas or water pressure, in the analysis. So naturally with the existence of more than one variable, not only must it be known how each variable varies with itself across the grid, but also how each variable varies with the other variables being considered.

The natural extensions of the variance and semivariogram to the multivariate case, which are known as the cross-covariance and the cross-covariograms, respectively, must now be considered.

$\bar{z}(x_i)=[z_1(x_i), z_2(x_i), \ldots, z_m(x_i)]$, which is the vector of values of each variable at each location. Once again, stationarity is assumed, so for a given random function $z_m(x)$

$E\{z_m(x)\}=\mu_m$, for all x, and

$C_{jk}(h)=E\{z_j(x+h)z_k(x)-\mu_j\mu_k$ is known as the cross covariance,

$2\gamma_{jk}(h)=E\{[z_j(x+h)-z_j(h)][z_k(x+h)-z_k(x)]\}$ is known as the cross-covariogram. Unlike the semivariogram, it can assume negative

values; this would be the case if the variables were negatively correlated. Another important difference is that the semivariogram is symmetric when j and k are reversed, but the cross-covariogram is not. At this point, a few other conditions that need to be pointed out are

(1) $\gamma_{jk}(h) = \gamma_{kj}(h)$

(2) $C_{jk}(h) = C_{kj}(-h)$

(3) $\gamma_{jk}(h) = \gamma_{jk}(-h)$

(4) $C_{jk}(h) \neq C_{jk}(-h)$.

The estimation procedure uses matrices:

$$[C_{jk}] = \begin{bmatrix} C_{11} & C_{12} & & C_{1n} \\ & \cdot & & \\ & \cdot & & \\ & \cdot & & \\ & \cdot & & \\ & \cdot & & \\ C_{m1} & C_{m2} & & C_{mn} \end{bmatrix}$$

The general equation for relating the cross semivariogram to the cross-covariance is

$2\gamma_{jk}(h)=2C_{jk}(0)-C_{jk}(h)-C_{kj}(h)$, and if $C_{jk}(h)=C_{kj}(h)$, then this will simplify to $\gamma_{jk}(h)= C_{jk}(0)-C_{jk}(h)$.

B. COKRIGING EQUATIONS

The idea is to estimate $\bar{z}(x)=[z_1(x), z_2(x), \ldots\ldots, z_m(x)]$ given the data $\bar{z}(x_1), \ldots\ldots, \bar{z}(x_n)$.

Here $\bar{z}^*(x)=\Sigma_i \bar{z}(x_i)\Gamma_i$, $i=1,\ldots,n$, and each Γ_i is an m by m matrix. This equation is equivalent to the one used in kriging except that the quantity being estimated is replaced by an mx1 vector, and the n weights are replaced by the n matrices Γ_i. It is sufficient for $\Sigma_i\Gamma_i = I_{nxn}$, $i=1,\ldots,n$, for $z^*(x)$ to be unbiased, or for

$\bar{z}_k^* = \Sigma_I\Sigma_j z_j(x_i)\lambda_{jk}^i$, where λ_{jk}^i= the entry in row j and column k of Γ_i, i.e. the weight given to variable j at location x_i in estimating variable k. Not each Γ_i is diagonal but $\Gamma_{jk}^i=0$, for $j \neq k$, and $\Sigma_i\lambda_{kk}^i = 1$ for all k.

The set of equations is(Hohn, pp.141-143):

$$\begin{bmatrix} \bar{C}^{11} & & \bar{C}^{1n} I \\ & \cdot & \\ & \cdot & \\ & \cdot & \\ \bar{C}^{n1} & & \bar{C}^{nn} I \\ I & & I\ 0 \end{bmatrix} \begin{bmatrix} \Gamma_1 \\ \cdot \\ \cdot \\ \cdot \\ \Gamma_n \\ \bar{\delta} \end{bmatrix} = \begin{bmatrix} \bar{C}^{10} \\ \cdot \\ \cdot \\ \cdot \\ \bar{C}^{n0} \\ I \end{bmatrix}$$

Each $\bar{C}^{i1}$ is an m by m matrix of cross covariances between sample i and sample 1. $\bar{\delta}$ is an m by m matrix of Lagrange multipliers. $\bar{C}^{i0}$ is an m by m matrix of cross covariances between each sample and the location to be sampled.

The estimation variance, denoted σ^2_E is equal to,

$$\sigma^2_E = \mathrm{Tr}[\ \bar{C}\ [0]] - \mathrm{Tr}[\Sigma_i\ \bar{C}_{i0}\Gamma],\ i=1,..,n\ -\ \mathrm{Tr}\ \ \bar{\delta}$$

$$= C_{kk}(0)\ -\Sigma_j\Sigma_i\ \bar{C}_{i0}\lambda_{jk}{}^{i}\ -\ \delta_{kk},$$

where i=1,....n, and j=1,.....,m, and Tr[A] is the trace of a matrix A. The variance is attributable to variable k.

$$\begin{bmatrix} \bar{\gamma}^{11} & & \bar{\gamma}^{1n}\ I \\ & \cdot & \\ & \cdot & \\ & \cdot & \\ & \cdot & \\ \bar{\gamma}^{n1} & & \bar{\gamma}^{nn}\ I \\ I & & I\ \ 0 \end{bmatrix} \begin{bmatrix} \Gamma_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \Gamma_n \\ \bar{\delta} \end{bmatrix} = \begin{bmatrix} \bar{\gamma}^{10} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \bar{\gamma}^{n0} \\ I \end{bmatrix}$$

Thus $\sigma^2_E = \mathrm{Tr}[\Sigma_i\ \bar{\gamma}^{0i}\Gamma_i]] + \mathrm{Tr}[\ \bar{\delta}\ ]$.

At this point it should be noted that $C_{jk}(h)=C_{kj}(h)$ is assumed throughout.

C. PRINCIPAL COMPONENT COKRIGING

Cokriging is more complex than ordinary kriging since there is more than one regionalized variable, and thus all the interactions between all of the regionalized variables must be accounted for. It can be very tedious and time consuming to compute and model n(n+1)/2 semivariograms(Hohn, pp. 165-167). In this section, a way of cokriging which does not require that the cross-semivariogram be modelled will be presented. In this method, the original variables are initially orthogonalized. The semivariograms are then computed and modelled. The transformed data are then kriged, and the estimates are reconstructed as a linear combination of the principal component estimates.

$$Z = \begin{bmatrix} z_{11} & z_{12} & & z_{1_n} \\ z_{21} & z_{22} & & z_{2_n} \\ & & \cdot & \\ & & \cdot & \\ & & \cdot & \\ & & \cdot & \\ z_{p1} & z_{p2} & & z_{pn} \end{bmatrix} \qquad [\hat{Z}] = \begin{bmatrix} \hat{z}_1 \\ \hat{z}_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \hat{z}_p \end{bmatrix} \qquad [1] = \begin{bmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix}$$

Here $\hat{z}_i = [\sum_j z_{ij}]/n$,

z is the data matrix,

n = the number of locations,

p = the number of variables,

[$\hat{z}$] is a px1 vector of variable means,

[1] is a px1 matrix of 1's, and

$$[S] = \begin{bmatrix} S_1 & 0 & & 0 \\ 0 & S_2 & & 0 \\ & & \cdot & \\ & & \cdot & \\ & & \cdot & \\ 0 & & 0 & S_p \end{bmatrix}$$

where [S] is a diagonal matrix of standard deviations S_i, and

$S_i = \sqrt{(\sum_j [(z_{ij} - \hat{z}_i)^2/(n-1)])}$, j = 1,2,....,n.

The data are then normalized to give a pxn matrix [χ]

$$[\chi] = [S]^{-1}([Z] - [\bar{z}][1]'),$$

where [1]' is the transpose of [1]. Each column of [χ] has mean of zero and standard deviation equal to one.

The matrix [R] of correlations between the variables is then calculated from the normalized data:

$$[R] = (1/n)[\chi][\chi]'.$$

The pxp matrix of eigenvectors [E] of [R] is then computed.

$$[E] = [E_1\ E_2\ \ldots\ E_p],$$

where each E_i is an eigenvector.

Define the matrices $[F]_{pxp}$ and $[A]_{pxn}$ as

$$[F] = [S][E]$$

and $[A] = [E]'[\chi]$.

The matrix [A] contains the coordinates of n samples along p axes. The weights in any row of [A] are uncorrelated with those in any other row of [A]. Should this absence of correlation extend spatially, i.e. for h>0, then the cross-semivariograms will resemble a pure nugget effect of zero.

The semivariograms are calculated from the rows of [A]. In order to verify that there exists negligible spatial covariance between the rows of [A], the cross-semivariograms must also be calculated and checked for orthogonality. For a given location, an estimate for variable k is calculated using ordinary kriging to give a_{k0}^*.

$$a_{k0}^* = \sum_i \lambda_i a_{ki}, \quad i=1,2,\ldots,n.$$

The p kriged estimates are assembled into a px1 vector.

$$[a_0^*] = \begin{vmatrix} a_{10}^* \\ a_{20}^* \\ \cdot \\ \cdot \\ \cdot \\ a_{p0}^* \end{vmatrix}$$

The matrix $[a_0^*]$ is then premultiplied by the transpose of [F] to transform and rescale back to the original variables.

$[z_n^*] = [\ \bar{z}\ ] + [F]'[a_0^*]$.

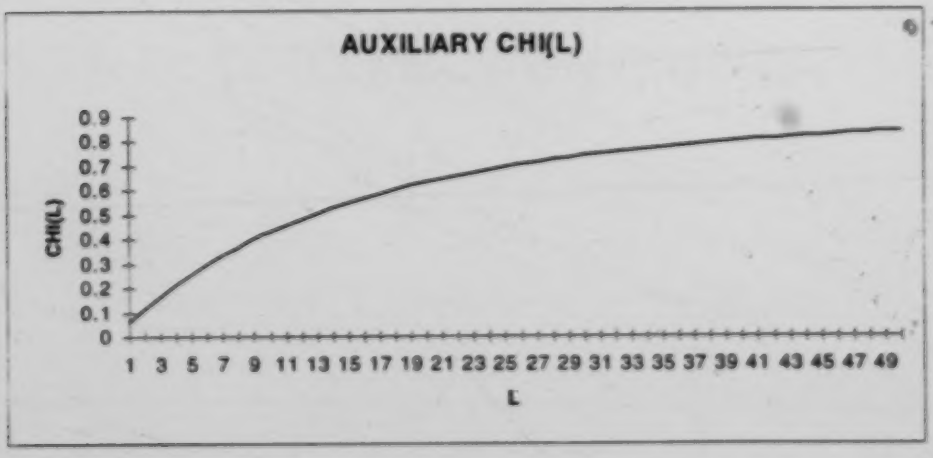
SECTION 14 EXAMPLE D:KRIGING AND AUXILIARY FUNCTION EXAMPLE(USING THE OKB2DM FORTRAN PROGRAM IN DEUTSCH)

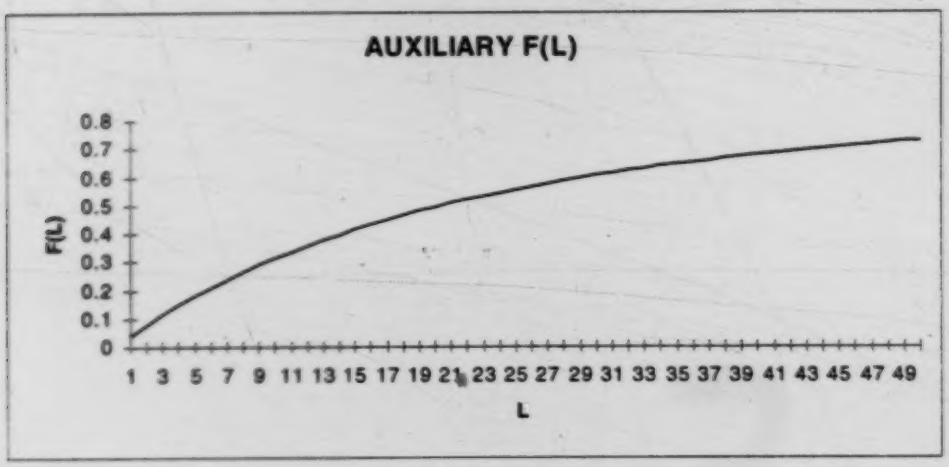
(Data on the IP's from reserves in West Virginia in APPENDIX D) For this example, the natural logarithms of the IP values were inputted into a computer program that performs the ordinary kriging calculation. It was a trial and error process in that a semivariogram model had to be chosen first, then the

parameters for the range, sill, and n had to be chosen. This process was continued until the kriging variance was minimized. The final model chosen was an exponential model with a=8, θ=0.7, and C=0.16. The kriging variances and the final kriged IP's of the unknown potential well-sites, along with their coordinates in the easting and northing direction are displayed in the following output:

easting	northing	ln(IP)	IP	variance	easting	northing	ln(IP)	IP	variance
499.16	4200	5.859	350.374	0.262	414.17	4320	5.559	259.56	0.262
400	4206.67	5.972	392.289	0.26	408.33	4326.67	5.593	268.54	0.263
400.83	4213.33	5.972	392.289	0.26	402.5	4333.33	5.624	277	0.263
400	4220	5.972	392.289	0.26	425	4340	5.604	271.51	0.259
400.83	4226.67	5.972	392.289	0.26	400.83	4346.67	5.711	302.17	0.261
400	4233.33	5.884	359.243	0.258	414.17	4353.33	5.604	271.51	0.259
400.83	4240	5.889	361.044	0.258	427.5	4360	5.61	273.14	0.258
400	4246.67	5.872	354.958	0.258	402.5	4366.67	5.604	271.51	0.259
413.33	4253.33	5.872	354.958	0.258	415.83	4373.33	5.609	272.87	0.258
407.5	4260	5.861	351.075	0.258	429.17	4380	5.637	280.62	0.258
401.67	4266.67	5.731	308.277	0.255	405	4386.67	5.609	272.87	0.258
400.83	4273.33	5.707	300.967	0.254	418.33	4393.33	5.647	283.44	0.258
401.67	4280	5.738	310.443	0.253	431.67	4400	5.646	283.16	0.258
448.33	4286.67	5.744	312.311	0.261	445	4406.67	5.647	283.44	0.259
405	4293.33	5.749	313.877	0.255	420.83	4413.33	5.643	282.31	0.257
403.33	4300	5.762	317.984	0.257	434.17	4420	5.632	279.22	0.258
416.67	4306.67	5.709	301.569	0.26	447.5	4426.67	5.571	262.7	0.258
400.83	4313.33	5.559	259.563	0.262					

The auxiliary functions for this example were then calculated using the parameters chosen for the exponential model. Here are the resulting graphs:





SECTION 15 DRIFT

A. BACKGROUND

A drift is a geographical complication that forces us to relax the assumption of second-order stationarity(Journel & Huijbregts, pp. 314-315). The following relationship occurs during the existence of drift:

$y(x) = z(x) - d(x)$, where $y(x)$ is the residual term, and $d(x)$ is the drift. It may or may not be stationary, but $E\{y(x)\}=0$, for all x.

The form of the drift $d(x)$ is assumed to be known.

It is assumed that $d(x) = \sum_l a_l * f_l(x)$, where $l = 1,.....,k$, where the a_l's are known constants and the $f_l(x)$'s are known functions.

B. TYPES OF DRIFT

(1)$d(x) = a_1 + a_2x$, linear.

(2)$d(x) = a_1 + a_2x + a_3x^2$, quadratic .

(3)$m(u,v)= a_1 + a_2u + a_3v + a_4u^2 + a_5u^2 + a_6uv$, the 2-dimensional case.

(4) $d(u,v,w,) = a_1 + a_2u + a_3v + a_4w + a_5u^2 + a_6v^2 + a_7w^2 + a_8uv + a_9uw + a_{10}vw + a_{11}u^3 + a_{12}v^3 + a_{13}w^3 + a_{14}uvw$.

But this, of course, may not be evident in some directions, i.e. some of the a_i may be zero.

(5)$d(x) = \ln(x)$, logarithmic drift.

(6)$d(x) = \exp(x)$, exponential drift.

Looking at the covariance or γ of a non-stationary random function, we have

$C(x,y) = E\{z(x)z(y)\} - \mu(x)\mu(y) = E\{Y(x)Y(y)\}$, and

$2\gamma(x,y)= E\{[z(x)-z(y)]^2\} - [\mu(x)- \mu(y)]^2 = Var\{Y(x) -Y(y)\} = E\{[Y(x) - Y(y)]^2\}$.

The semivariogram $\gamma(x,y)$ represents the semivariogram of the residuals $Y(x) = Z(x) - \mu(x)$. Due to the existence of drift, $\gamma(x,y)$ cannot be estimated directly from the initial experimental data. This is so because in order to do so, one would have to simultaneously estimate the drift $d(x)$ and the semivariogram $\gamma(x,y)$ from a single realization $z(x)$. This is not possible, although some iterative methods have been proposed.

SECTION 16 NONLINEAR GEOSTATISTICS

If a drift exists but the covariance or semivariogram and the two or k-variate

distributions of the random function z(x) are known, then nonlinear geostatistics must be used(Journel & Huijbregts, pp.555-559).

SECTION 17 A VECTORIAL SPACE AND PROJECTION MODEL OF KRIGING (Journel & Huijbregts, pp. 558-559)

Define $A=\{Z(x), x\in D\}$ to be a set of random variables Z(x) defined at each point x of, say a tridimensional deposit. Let E be the vector space of all finite linear combinations of the elements of A plus the limits of all such finite linear combinations(this is known as the closure property), i.e. $E=\{\sum_\beta \lambda_\beta Z(x_\beta);\ Z(x_\beta)\in A,\ \lambda_\beta \text{ real}\}$.

E is equipped with a scalar product equal to the non-centred covariance(which is not necessarily stationary), given by

$$\langle Z(x),Z(y)\rangle = E\{Z(x)Z(y)\} = \sigma_{xy}.$$

Let $Z(x_0)=Z_0\in E$ be any unknown variable and E' be any vector subspace of E(or in general, any closed linear manifold). According to the Projection Theorem, there is one and only one element $Z^*\in E'$ which will minimize the distance $\|Z_0-Z^*\|$; this unique element is known as the projection of Z_0 onto the subspace E'.

Kriging is simply the projection of some unknown value onto a particular subspace E', within which an estimator Z^* is searched for. The corresponding minimum estimation variance, $\|Z_0-Z^*\|^2=E\{[Z_0-Z^*]^2\}$, is known as the kriging variance. There are just as many kriging processes, and thus kriging estimators Z^*, as there are different

subspaces E'⊂E within which the projection of the unknown element $Z(x_0)$ is carried out. Consider two subspaces E" and E' of E, such that E"⊂E'⊂E. Their corresponding kriging estimators are therefore Z*" and Z*', respectively. Since E" is contained in E', the projection Z*' will be nearer to the unknown than Z*". In terms of estimation variance, this means that $\|Z_0-Z^{*'}\|^2 \leq \|Z_0-Z^{*''}\|^2$. Therefore as the set where the search for the estimator takes place gets larger, the estimation improves.

Projection theorem: There exists one and only one element $z^* \in E'$ which will minimize $\|z_0-z^*\|$; this unique element is the projection of z_0 onto E' and the kriging variance is $\|z_0 - z^*\|^2$

$E'' \in E' \in E$, and generalizing this gives

$E^{(n)} \in E^{(n-1)} \in \ldots\ldots\ldots\ldots, \in E'' \in E' \in E$

Another possible method of finding z* is to use the generalization of Heren's formula to 3-space. In 2-space, Heren's formula states that the distance from a point (h,k) to a line with equation ax + by = c , is

$$| (ah + bk + c)/ \sqrt{a^2+b^2} |$$

If this is generalized to the 3 space case, then the distance from a point (h,k,l) in space to a plane ax + by + cz = d , is

$|(ah + bk + cl + d)/\sqrt{a^2+b^2+c^2}\ |$.

The coordinates of z^* could then be obtained from this formula.

SECTION 18 LINEAR KRIGING PROCESSES

A. BACKGROUND

The vector subspace $E^{(n+1)} \subset E$ of dimension n+1 is generated by the linear combinations:

$\Sigma_\alpha \lambda_\alpha z_\alpha + \lambda_0 * 1$ of n particular variables $\{z_\alpha = z(x_\alpha);\ \alpha = 1,,,,n\}$, called data plus the constant 1 (Journel & Huijbregts, pp.559-567). Linear kriging processes project the unknown $z(x_0)$ onto $E^{(n+1)}$ itself or any linear manifold of $E^{(n+1)}$.

Consider the expectation of an element $z^* = \lambda_0 + \Sigma_\alpha \lambda_\alpha z_\alpha \in E^{(n+1)}$, we have

$E[z^*] = \lambda_0 + \Sigma_\alpha \lambda_\alpha E[z_\alpha] = \lambda_0 + \Sigma_\alpha \lambda_\alpha \mu(x_\alpha)$.

The estimator $z^*(x)$ of $z_0(x)$ is unbiased if and only if

$\lambda_0 + \Sigma_\alpha \lambda_\alpha \mu(x_\alpha) = \mu(x_0)$ {**}

B. CASES

(1) All expectations are known(stationary or not). In this case $\lambda_0 = \mu(x_0) - \Sigma_\alpha \lambda_\alpha \mu(x_\alpha)$, therefore an unbiased estimator exists.

(2) The expectation of z(x) is stationary but unknown, i.e.

$E\{z(x)\} = \mu$ for all x if $\lambda_0 = \mu - \mu[\Sigma_\alpha \lambda_\alpha] = \mu[1 - \Sigma_\alpha \lambda_\alpha]$, therefore it is unbiased

(3) The expectation $\mu(x)$ is neither stationary nor known. In this case the unbiasedness relation {**} cannot be expressed. In this case, the form of the expectation $\mu(x)$ must be provided. For example if $\mu(x)$ is an unknown linear combination of L known functions f(x),

then $\mu(x) = \Sigma_l [a_l f_l(x)]$, with all a_l's unknown. The unbiasedness condition {**} can now be rewritten as:

$\lambda_0 + \Sigma_l a_l \Sigma_\alpha \lambda_\alpha f_l(x_\alpha) = \Sigma_l a_l f_l(x_0)$.

Regardless of the unknown parameters, this relation will be satisfied if and only if $\lambda_0 = 0$,

which results in the restriction $E^n \subset E^{n+1}$ and

$$\Sigma_\alpha \lambda_\alpha f_l(x_\alpha) = f_l(x_0), \text{ for } l = 1 \text{ to } L.$$

Case (1) Known Expectation:

Let $z^*_{k0} = \lambda_0 + \Sigma_\alpha \lambda_{0\alpha} z_\alpha$ be the projection of the unknown $z(x_0)$ onto $E^{(n+1)}$. It is unique and defined by $z(x_0)-z^*_{k0}$ as being orthogonal to each of the n+1 vectors that generate $E^{(n+1)}$, i.e. $\langle z(x_0)- z^*_{k0}, 1\rangle = 0$, and

$\langle z(x_0)- z^*_{k0}, z_\alpha\rangle = 0$, for all $\alpha = 1,,,,,,,n$

The kriging variance is again

$\|z(x_0) - z^*_{k0}\|^2$

Heren's formula could be used again to find z^*_{k0}.

Case(2) Unknown stationary expectation

Let $C_1 \subset E^n \subset E^{(n+1)}$. In order for {**} to be satisfied, the following conditions must be imposed.

(1) $\lambda_0 = 0$, and

(2) $\Sigma_\alpha \lambda_\alpha = 1$.

The first condition is required to restrict the set of possible estimators to the vector subspace E^n. These vectors are generated by the linear combinations $\Sigma_\alpha z_\alpha$ of the n data only. The second condition is required to restrict E^n to the linear manifold C_1. Then $z^*_{k1} = \Sigma_\alpha \lambda_{k1\alpha} z_\alpha$ is unique and is defined by

(a) $z^*_{k1} \in C_1$, i.e. $\Sigma_\alpha \lambda_{k1\alpha} = 1$

(b) $z(x_0) - z^*_{k1}$ is orthogonal to each of the n vectors $z_\alpha - y$

generating C_1 with $z_\alpha, y \in C_1$

$\langle z(x_0) - z^*_{k1}, z_\alpha - y \rangle = 0$ for all $\alpha = 1,,,,,n$.

This gives us a system of n+1 linear equations

$\Sigma_\beta \lambda_{k\beta} = 1$

$\Sigma_\beta \lambda_{k\beta} \sigma_{\alpha\beta} - \mu_1 = \sigma_\alpha x_0$, for all $\alpha = 1,,,,,,,,n$

Once again, Heren's formula could be used.

Case(3) Linear Kriging in the existence of a drift.

The expectation of z(x) is non-stationary and unknown, but is of the following known form

$E\{z(x)\} = \mu(x) = \Sigma_l a_l f_l(x)$, l=1,.....,L

The unbiasedness condition forces us to restrict our search for a linear estimator to the linear manifold $C_L \subset E^n$ defined by the following condition on the weights

$\Sigma_\alpha \lambda_\alpha f_l(x_\alpha) = f_l(x_0)$, for all l=1,...,L,

The kriging estimator $z^*_{kL} = \Sigma_\alpha \lambda_{kL\alpha} z_\alpha$ is the projection of unknown $z(x_0)$ onto C_L.

Therefore we have (n+L) linear equations

$\Sigma_\beta \lambda_{kL\beta} f_l(\beta) = f_l(x_0)$, for all l= 1 to L,

$\Sigma_\beta \lambda_{k\beta} \sigma_{\alpha\beta} - \Sigma_l \mu_l^* f_l(x_\alpha) = \sigma_\alpha x_0$, for all l= 1,,n

Once again Heren's formula could be used.

SECTION 19 ESTIMATION VARIANCE

Consider the case where the petroleum geologist wants to estimate the IP's in a block W from the values of n samples w_i(i= 1,,,,n), which may be located inside or outside of W(Rendu, pp. 54-56).

Let w_s be the set of all samples w_i.

$w_s = [w_1;w_2;...;w_n]$

and the block W is estimated using the samples w_s.

Let x_i = the value of the i^{th} sample , and

μ_w = the average value of the samples in block W.

Let $\hat{\mu}_w$ = the estimator of μ_w, where

$$\hat{\mu}_w = [\Sigma_i w_i x_i]/[\Sigma_i w_i],\ i=1,.....,n$$

The mean squared error when estimating μ_w by $\hat{\mu}_w$ is known as the estimation variance of W by w_s. It is the error incurred when assigning the value of sample set w_s to the block W, and is denoted $\sigma^2_E(w_s \text{ to } W)$.

By definition, $\sigma^2_E(w_s \text{ to } W) = E[(\hat{\mu}_w - \mu_w)^2]$

If n=1, the expression for $\sigma^2_E(w_s \text{ to } W)$ is known as the extension variance. The estimator

$\hat{\mu}_w$ is unbiased because the weights assigned to the sample values all sum up to 1.

Therefore $E[(\hat{\mu}_w - \mu_w)] = 0$.

The error term $(\hat{\mu}_w - \mu_w)$ is assumed to be normally distributed with zero mean and variance equal to $\sigma^2_E(w_s \text{ to } W)$. As a result, confidence limits for μ_w can be obtained.

Let t_p = the value of the normally distributed variable t such that $P[t < t_p] = p$.

Then $P[\hat{\mu}_w < \mu_w - t_{1-p}\, \sigma^2_E(w_s \text{ to } W)] = p$, and

$P[\hat{\mu}_w > \mu_w + t_{1-p}\, \sigma^2_E(w_s \text{ to } W)] = p$.

An expression for $\sigma^2_E(w_s \text{ to } W)$ will now be derived.

Consider the sample set w_s as being made up of p points with value $x_j (j=1,2,\ldots,p)$ and the block W consists of q points with value $x_k (k=1,2,\ldots,q)$. The terms x_j and x_j' will be used to denote two points in w_s ($j'=1,2,\ldots,p$, and $x_j' = x_j$ if $j'=j$), and the terms x_k and x_k' will denote two points in W($k'=1,2,\ldots,q$, and $x_k' = x_k$ if $k'=k$).

Therefore we can write

$$\sigma^2_E (w_s \text{ to } W) = E[(\hat{\mu}_w - \mu_w)^2]$$

$$= E[((1/p)\sum_i x_i - (1/q)\sum_k x_k)^2]$$

$$= -(1/p^2)\sum_i \sum_{i'} E[(1/2)(x_i - x_{i'})^2] - (1/q^2)\sum_k \sum_{k'} E[(1/2)(x_k - x_{k'}) + (1/pq)\sum_i \sum_k E[(x_i - x_k)^2]$$

$$= - \bar{\gamma} (w_s;w_s) - \bar{\gamma} (W;W) + 2 \bar{\gamma} (w_s;W).$$

Thus we now have the fundamental relationship:

$$\sigma^2_E (w_s \text{ to } W) = - \bar{\gamma} (w_s;w_s) - \bar{\gamma} (W;W) + 2 \bar{\gamma} (w_s;W).$$

Or if we wish to use the covariogram

$$\sigma^2_E (w_s \text{ to } W) = \bar{c} (w_s;w_s) + \bar{c} (W;W) - 2 \bar{c} (w_s;W).$$

SECTION 20 RANDOM KRIGING

Often in practice, the samples being used to estimate a particular block are not located on a regular grid(Rendu, pp.68-70). The sample sizes may vary, and the number and relative position of the samples used to estimate each block vary from block to block. Thus if each sample in the neighbourhood of a block W were examined individually, estimation of W would be very tedious. A possible solution is to group the samples in blocks W_i(i=1,,,,,n) in the neighbourhood of W, and then estimate μ_W as a weighted average of the mean x_i of the sampled blocks in the W_i. In order to calculate the kriging estimator of μ_W one must be able to calculate the variance of x_i, the

covariance of x_i and x_j, and the covariance of x_i and μ_w . The samples are assumed

to be randomly distributed in W_i. Kriging with this assumption is therefore known as random kriging.

The following notations are used

W = the block being valued.

n = the number of blocks used in the valuation of W.

W_i = the i^{th} block used in the valuation of W(i=1,2,....,n).

w = size of a unit sample(it is assumed that all samples are of size w).

q_i = the number of samples of size w in W_i.

x_{ij} = the value of the j^{th} sample of size w in block W_i(j=1,2,....,q_i).

x_i = the average value of x_{ij} in block W_i.

w_{ij} = the support of x_{ij} .

w_i = the support of x_i (w_i is simply the union of all the w_{ij}).

or $w_i = [w_{i1};w_{i2};...;w_{iq}]$

μ_w = the unknown value of block W

μ_x = the kriging estimator of μ_w

μ_x = $\sum b_i x_i$, where the b_i terms satisfy the kriging equations.

This system will therefore be a function of

$\bar{\gamma}(w_i;w_i)$, i=1,2,...,n

$\bar{\gamma}(w_i;w_j)$, i=1,2,...,n;j=1,2,...,n.

If the positions of the q_i samples of size w in W_i are unknown, then the values of $\bar{\gamma}$ cannot be calculated. If, however, the q_i samples take all possible positions in W_i at random, then the expected value of $\bar{\gamma}$ can be calculated. These expected values are used in the kriging systems of equations. From the definition of $\bar{\gamma}$, if the samples w_q are randomly distributed in W_a, then the expected value of $\bar{\gamma}(w_i;W)$ is

$E[\bar{\gamma}(w_i;W)] = \bar{\gamma}(W_i;W)$ i=1,,,,n

$E[\bar{\gamma}(w_i;w_j)] = \bar{\gamma}(W_i;W_j)$ for $i \neq j$

$\bar{\gamma}(w_i;w_i) = (1/q_i^2)\Sigma_j \bar{\gamma}(W_q;W_q) + 2\Sigma_k\Sigma_j \bar{\gamma}(w_q;w_a)$

(where j=1,.....,q_i for the first expression and for the second expression j=1,.....,q_{k-1} and k=j+1,.....,q_k, since all the w_q's have the same size.

$E[\bar{\gamma}(w_q;w_a)] = (W_i\ \bar{\gamma}(W_i;W_i) - w\ \bar{\gamma}(w;w))/(W_i - w)$, for $j \neq k$.

Random kriging is only valid if the block size W_i is much larger than the sample size w.

In the limiting case as w approaches W_i,

$$\frac{\partial}{\partial wi}\{[W_i\ \bar{\gamma}\ (W_i;W_j) - w\ \bar{\gamma}\ (w;w)]/(W_i-w)\} = \bar{\gamma}\ (w;w) + w\ \bar{\gamma}\ '(w;w).$$

where $\bar{\gamma}\ '(w;w)$ = the derivative of $\bar{\gamma}\ (w;w)$.

Therefore the limit as w goes to W_i of this derivative is

$\lim_{w \to W_i}\{\bar{\gamma}\ (w;w) + w\ \bar{\gamma}\ '(w;w)\}$

$= \bar{\gamma}\ (W_i;W_j) + W_i\ \bar{\gamma}\ '(W_i;W_j)$.

Thus $E[\bar{\gamma}\ (w_j;w_k)] = \bar{\gamma}\ (W_i,W_j)$ for $j \neq k$,

and $E[\bar{\gamma}\ (w_j;w_j)] = (1/q_j)\ \bar{\gamma}\ (w;w) + ((q_j-1)/q_j)\ \bar{\gamma}\ (W_i;W_j)$,

and so in the limiting case as w goes to W_i, this becomes

$(1/q_j)\ \bar{\gamma}\ (W_i;W_j) + ((q_j - 1)/q_j)\ \bar{\gamma}\ (W_i;W_j) = \bar{\gamma}\ (w_j;w_j)$.

SECTION 21 DISPERSION VARIANCE

Consider a large block W(x) in Ω with average value $\mu_{W(x)}$, and all possible smaller blocks or samples w(x') of size w in W(x). The variance of the blocks w(x') in the block W(x) is known as the dispersion variance(Rendu, pp. 45-48). It is determined from the following relationship:$\sigma^2[w \text{ in } W(x)]=E_{w(x')\text{in } W(x)}\{[\mu_w(x')-\mu_W(x')]^2\}$. If stationarity is present, this variance is a function only of the dimensions of the supports w and W, and is independent of the position of the block W(x). The dispersion variance of w in W can therefore be defined as: $\sigma^2(w \text{ in } W)=E_{w \text{ in } W}[(\mu_w-\mu_W)^2]$. If stationarity does not exist, the variance of samples w in blocks W may vary only slightly from block to block. In this case the average dispersion variance of w in W(x) would be used for all possible W(x) in Ω: $\sigma^2(w \text{ in } W)=E_{W(x) \text{ in } \Omega}\{\sigma^2[w \text{ in } W(x)]\}$.

The variance of $\mu_w(x)$ when w(x) assumes all positions in Ω is called the dispersion variance of w in Ω and is denoted by

$\sigma^2(w \in \Omega) = E_{[w(x) \in \Omega]}\{[\mu(x) - \mu]^2\}$.

Let w, W' and W" are 3 blocks of increasing size

The variance additivity relationship is defined in the following way

$\sigma^2(w \in W") = \sigma^2(w \in W') + \sigma^2(W' \in W")$.

The generalized formula is

$\sigma^2(w \in W^n) = \sigma^2(w \in W') + \sigma^2(W' \in W'') + \ldots\ldots\ldots\ldots + \sigma^2(W^{(n-1)'} \in W^n)$.

To find any $\sigma^2(W_i' \in W_i'')$, the above formula is applied.

From this equation. The following can be deduced

$\sigma^2(w \in W^n) \geq \sigma^2(W^{(n-1)'} \in W^n)$

SECTION 22 CONDITIONAL SIMULATIONS

An infinite number of possible realizations $\{z_s(x), s=1,\ldots,\infty\}$ of a regionalized variable z(x) exist (Journel & Huijbregts, pp. 492-496). From these, the simulations $z_{sc}(x)$ that are chosen are the ones that meet the experimental data values at the actual data locations x_α, i.e. the simulations for which $z_{sc}(x_\alpha)=z_0(x_\alpha)$, for all x_α. This process is called conditioning the simulation to the experimental data, or conditional simulation. The simulated wells and the real wells have the same clusters of rich and poor data at each location. The regionalization of $z_0(x)$ is then considered. This is the realization of a stationary random function $z_0(x)$ with expectation μ, covariance C(h), and semivariogram $\gamma(h)$. The conditionally simulated process must pass through the data and will have the same expectation and second order moments C(h) or $\gamma(h)$, i.e. $z_{sc}(x_\alpha) = z_0(x_\alpha)$, for all x_α's that are elements of the data set(Cressie, pp 207-208).

WHY SIMULATIONS ?

The main objective of estimation is to provide an estimator $z^*(k)$ at each point x which is as close as possible to $z_0(k)$, the true unknown IP. As mentioned before, the criteria for measuring the usefulness of estimation are unbiasedness and estimation variance. These estimators, however, need not reproduce the spatial variability of the true *IP's* $\{z_0(k)\}$. When minimizing the estimation variance in kriging, this involves a smoothing of the true dispersions(Cressie, pp 208). As a result, the variance of the kriging estimate underestimates the local variability of the true IP's. Conditional simulation, on the other hand, has the same first two experimentally found moments(mean and covariance) as the real IP's $\{z_0(k)\}$. It therefore is better able to identify uncertainty of the true IP's. But on the contrary, the conditionally simulated value, $z_{sc}(x)$, is not the best possible estimator of $z_0(x)$.

Generally speaking, the objectives of simulation and estimation address two different problems. Although estimation in general is closer on average to the true values, simulations provide a better reproduction of the fluctuations of the real curve. Estimation is preferred if the objective is to locate and estimate reserves, while simulation is preferred if the objective is to study the dispersion of the characteristics of the IP's. As well, one must keep in mind that in practice, the real curve is known only at the experimental data points.

$z_0(x)$[true value] and $z^*_{ok}(x)$[kriged value] differ by an unknown error

$z_0(x) = z^*_{ok}(x) + [z_0(x) - z^*_{ok}(x)] =$

To get the desired conditional simulation, it suffices to replace the unknown kriging

error $[z_0(x) - z^*_{ok}(x)]$ by an independent isomorphic kriging error $[z_s(x) - z^*_{sk}(x)]$. Given a realization $z_s(x)$, the kriging procedure will give a kriging error $[z_s(x) - z^*_{sk}(x)]$ which is isomorphic to the true error $[z_0(x) - z^*_{ok}(x)]$ and independent of $z^*_{ok}(x)$ when applied to the simulated data. The desired conditional simulation is then written as

$$Z^*_{sc}(x) = z^*_{ok}(x) + [z_s(x) - z^*_{sk}(x)]$$

The conditioning principle gives both values kriged $z^*_{ok}(x)$ and $z^*_{sc}(x)$ at each point x by considering the independence of E(x) and $[z_s(x) - z^*_{sk}(x)]$. The variance of the estimation of the real $z_0(x)$ by the conditional simulation $z_{sc}(x)$ is written as

$$E\{[z_0(x)-z_{sc}(x)]^2\}=E\{[z_0(x)-z^*_{ok}(x)]^2\} + E\{[z_s(x)-z^*_{sk}(x)]^2\}$$

$$=2E\{[z_0(x)-z^*_{ok}(x)]^2\} = 2\sigma_k.$$

This result is not surprising given the greater variability in conditional simulations.

SECTION 23 REGULARIZATION

If the values of z(x) cannot be measured at a particular point x on the grid, one must use what is called regularization. The average of the points surrounding x is calculated, and this average w_i is taken to be the value of z at x. Therefore an area Ω surrounding x must first be analysed. The measured property w_i is referred to as a regularized variable(Rendu, pp. 22-24), and the corresponding semivariogram or

covariogram computed from these values is also referred to as regularized.

Fundamental relationships:

It is assumed that Ω has been investigated by samples of equal size w, and therefore the semivariogram or covariogram computed using these sample values is regularized by the volume or the area w.

Notation:

w_1 = the centre of mass at x,

w_2 = the centre of mass at x+h,

$z(x)$ = the value of the point regionalized variable z at point x,

μ = average value of z in Ω,

$\gamma(h)$ = the point semivariogram of $z(x)$,

$\sigma(h)$ = point covariogram of $z(x)$,

σ^2 = variance of z in Ω,

$w_1(x)$ = sample of size w_1 centred at point x,

$w_2(x+h)$ = sample of size w_2 centred at point x+h,

$\mu_{w1}(x)$ = average value of z in $w_1(x)$,

$\gamma_{w1}(h)$ = the regularized semivariogram of $\mu_{w1}(x)$,

$\sigma_{w1}(h)$ = the regularized covariogram of $\mu_{w1}(x)$,

σ^2_{w1}= variance of $\mu_{w1}(x)$ in Ω.

$\bar{\gamma}(w_1,w_2)$ is defined as the average value of the point semivariogram $\gamma(x'x'')$ where

x'x" is the vector distance between any two points x' in the sample $w_1(x)$ and x" in the sample $w_2(x+h)$. This integral is a function of the sample shape w and of h, the distance between the samples $w_1(x)$ and $w_2(x+h)$. It is independent of the position of the samples in Ω and of the sample values $\mu_{w1}(x)$ and $\mu_{w1}(x+h)$.

$\bar{\gamma}\ (w_1,w_2) = (1/w_1w_2) \quad \int dx' \int dx'' \gamma(x'x'')\{x' \text{ in } w_1,\ x'' \text{ in } w_2\}$

$\bar{\gamma}\ (w_1;w_1)$ is equal to the average value of $\gamma(x'x'')$ where x' and x" are any two points in the sample w_1.

$\bar{\sigma}\ (w_1;w_2)$ is equal to the average value of $\sigma(x'x'')$ where x' and x" are any two points in $w_1(x)$ and $w_2(x+h)$ respectively.

$\bar{\sigma}\ (w_1;w_1)$ is equal to the average value of $\sigma(x'x'')$, where x' and x" are any two points in the sample w_1 .

By definition, the following relationships are true:

$\mu = E_{x \text{ in } \Omega}[z(x)]$,

$\mu_{w1}(x) = E_{x' \text{ in } w1(x)}[z(x')]$,

$\sigma(h) = E_{x \text{ in } \Omega, x+h \text{ in } \Omega}\{[z(x)-\mu][z(x+h)-\mu]\}$,

$\sigma_{w1}(h) = E_{w1(x) \text{ in } \Omega, w2(x+h) \text{ in } \Omega}\{[\mu_{w1}(x)-\mu][\mu_{w2}(x+h)-\mu]\}$,

$2\gamma(h) = E_{x \text{ in } \Omega, x+h \text{ in } \Omega}\{[z(x)-z(x+h)]^2\}$,

$2\gamma_{w1}(h) = E_{w1(x) \text{ in } \Omega, w2(x+h) \text{ in } \Omega}\{[\mu_{w1}(x)-\mu_{w2}(x+h)]^2\}$.

The following relationships will now be proved:

(1) $\gamma_{w1}(h) = \bar{\gamma}(w_1;w_2) - \bar{\gamma}(w_1;w_1)$.

(2) $\sigma_{w1}(h) = \bar{\sigma}(w_1;w_2)$.

To prove (2), assume that $w_1(x)$ is made up of n points x_i', {i=1,,,n} with value $z_i'=z(x_i')$,

and that $w_2(x)$ is made up of m points x_j'', {j=1,,,m} with value $z_j'' =z(x_j'')$

therefore $\mu_{w1}(z)=(1/n)\Sigma_i z_i'$, i= 1,.....,n

and $\mu_{w2}(z+h)=(1/m)\Sigma_j z_j''$, j= 1,.....,m

So $\sigma_{w1}(h)=E\{[(1/n)\Sigma_i z_i'- \mu][(1/m)\Sigma_j z_j'' - \mu]\}$

$=E\{(1/nm)[\Sigma_i z_i'-\mu][\Sigma_j z_j'' - \mu]\}$

$=(1/nm)[\Sigma_i\Sigma_j[E[(z(x_i')- \mu)(z(x_j'')- \mu)]]$

$=(1/nm)\Sigma_i\Sigma_j[\sigma(x_i'x_j'')]$

$= \bar{\sigma}(w_1;w_2)$. QED for (2)

To prove (1),

We first consider that $\gamma(h) = \sigma^2 - \sigma(h)$,

then $\gamma_{w1}(h) = \sigma^2_w - \sigma_{w(h)}$,

and since $\sigma^2_{w1}= \gamma_{w1}(0) = \bar{\sigma}(w_1;w_1)$

this implies $\gamma_{w1}(h) = \delta(w_1;w_1) - \delta(w_1;w_2)$

which implies $\delta(w_1;w_1) = \sigma^2 - \bar{\gamma}(w_1;w_1)$

and since $\bar{\gamma}(w_1;w_2) = \sigma^2 - \bar{\gamma}(w_1;w_2)$,

we get $\gamma_{w1}(h) = \bar{\gamma}(w_1;w_2) - \bar{\gamma}(w_1;w_1)$. QED for (1)

SECTION 24 CONCLUSION

The purpose of this thesis has therefore been accomplished. The original geostatistical concepts have been explained and represented mathematically. The original assumptions were laid out in order to deal with the simplest cases, then some of these were relaxed and other models were proposed to deal with the more complicated cases. In some cases, by adding new parameters, further modification allowed greater latitude in fitting the models to the data. As well, we were able to show in some cases that by allowing a limiting case, we were better able to represent some real world cases. As a result of the generalizations, we were able to better fit the models to the data.

SECTION 25 APPENDIX A: DATA FOR FITTING BASIC SEMIVARIOGRAMS

The notation used is:

h = distance,

gamma(h) = the experimental semivariogram values ,

sph(a,φ) = a fitted spherical semivariogram model with range a and spherical parameter φ,

exp(1) = the fitted regular exponential model, and

exp(a,θ) = a fitted exponential semivariogram model with range a and exponential parameter θ.

For each value of h, the corresponding values for each of the terms above is in the corresponding column on the same line.

h	gamma(h)	sph(1.6,3.5)	exp(1)	exp(.63,.72)
0		0.078		
0.18	0.115	0.094	0.094	0.111
0.56	0.137	0.126	0.121	0.138
1.02	0.152	0.159	0.142	0.154
1.51	0.163	0.177	0.156	0.163
2	0.165	0.178	0.164	0.168
2.51	0.168	0.178	0.170	0.171
3	0.174	0.178	0.173	0.173
3.5	0.172	0.178	0.175	0.175
4	0.172	0.178	0.176	0.176
4.5	0.175	0.178	0.177	0.176
5	0.176	0.178	0.177	0.177
5.49	0.178	0.178	0.178	0.177
6	0.175	0.178	0.178	0.177

6.5	0.177	0.178	0.178	0.178
7	0.177	0.178	0.178	0.178
7.5	0.181	0.178	0.178	0.178
8	0.179	0.178	0.178	0.178
8.5	0.176	0.178	0.178	0.178
9	0.179	0.178	0.178	0.178
9.5	0.176	0.178	0.178	0.178
10	0.173	0.178	0.178	0.178
10.5	0.174	0.178	0.178	0.178
11	0.175	0.178	0.178	0.178
11.51	0.178	0.178	0.178	0.178
12	0.178	0.178	0.178	0.178

SECTION 26 APPENDIX B: DATA FOR FITTING A DOUBLE SPHERICAL MODEL

Here we have

h = distance,

gamma(h) = the experimental semivariogram values,

sph(a_1,φ_1:a_2,φ_2) = a fitted double spherical semivariogram with ranges a_1 and a_2, and spherical parameters φ_1 and φ_2.

h	gamma(h)	sph(1,2.5:12,3.3)
0.39	6.140	5.640
1.04	8.318	7.238
2.03	9.411	8.415
3.02	9.205	9.565
4.02	10.760	10.689
5.02	10.870	11.757
6	12.160	12.732
7	11.890	13.635
8	14.320	14.425
9	17.570	15.078
10.01	17.110	15.578
10.99	19.750	15.887
12	20.080	16.000
13	23.220	16.000
14	20.100	16.000
14.99	22.000	16.000
16.01	20.560	16.000

17.02	17.620	16.000
18	17.110	16.000
19.01	15.790	16.000
19.99	16.660	16.000
20.99	15.070	16.000
22	15.120	16.000
22.99	17.730	16.000

SECTION 27 APPENDIX C: DATA FOR FITTING AN ANISOTROPIC MODEL

In this appendix, we have

h = distance,

actual = the experimental semivariogram,

theoretical = the theoretical semivariogram model fitted(could be any kind),

b-0 = the experimental semivariogram in the east-west direction, i.e. 0 degrees,

c-45 = the experimental semivariogram in the northeast-southwest direction, i.e. 45 degrees,

d-90 = the experimental semivariogram in the north-south direction, i.e. 90 degrees,

e-135 = the experimental semivariogram in the northwest-southeast direction, i.e. 135 degrees,

gamma(e-w) = the theoretical semivariogram model fitted in the east-west direction, i.e. 0 degrees,

gamma(ne-sw) = the theoretical semivariogram model fitted for the northeast-southwest direction, i.e. 45 degrees,

gamma(n-s) = the theoretical semivariogram model fitted for the north-south direction, i.e. 90 degrees,

gamma(nw-se) = the theoretical semivariogram model fitted for the northwest-southeast, i.e. 135 degrees,

h_u' = the coordinate h_u' ,

h_v' = the coordinate h_v' .

EAST-WEST DIRECTION

h	actual b-0	theoretical gamma(e-w)
0		
0.18	40	103.656
0.56	95	111.375
1.02	115	120.718
1.51	118	130.664
2	133	140.593
2.51	142	150.886
3	144	160.696
3.5	149	170.573
4	168	180.234
4.5	180	189.576
5	195	198.463
5.49	197	206.569
6	207	214.163
6.5	195	220.523
7	224	225.518
7.5	212	228.807
8	227	230.000
8.5	242	230.000
9	205	230.000
9.5	233	230.000
10	223	230.000
10.5	222	230.000
11	231	230.000
11.51	229	230.000
12	235	230.000
12.5	226	230.000
13	239	230.000
13.5	243	230.000
14	217	230.000
14.5	227	230.000
15	248	230.000
15.5	232	230.000
16	234	230.000
16.5	243	230.000
17	238	230.000
17.5	241	230.000
18	229	230.000
18.5	250	230.000

NORTH-SOUTH DIRECTION

h	actual d-90	theoretical gamma(n-s)
0		
0.18	105	106.169
0.56	80	115.651
1.02	121	125.140
1.51	127	133.885
2	134	141.636
2.51	138	148.886
3	148	155.207
3.5	141	161.106
4	162	166.523
4.5	172	171.519
5	159	176.141
5.49	166	180.344
6	172	184.411
6.5	173	188.122
7	178	191.584
7.5	189	194.818
8	198	197.844
8.5	207	200.677
9	209	203.334
9.5	204	205.828
10	204	208.170
10.5	233	210.372
11	218	212.443
11.51	223	214.431
12	229	216.231
12.5	253	217.963
13	238	219.596
13.5	245	221.138
14	238	222.593
14.5	245	223.968
15	241	225.267
15.5	229	226.496
16	247	227.657
16.5	259	228.757
17	276	229.798
17.5	255	230.783
18	255	231.717
18.5	259	232.601

19	259	230.000	19	261	233.440

NORTHEAST-SOUTHWEST DIRECTION

h	theoretical gamma(ne-sw)	actual c-45
0		
0.18	105.297	104
0.56	112.973	79
1.02	120.696	120
1.51	127.984	126
2	134.644	133
2.51	141.086	137
3	146.902	147
3.5	152.524	140
4	157.876	161
4.5	162.993	171
5	167.900	158
5.49	172.526	165
6	177.167	171
6.5	181.558	172
7	185.804	177
7.5	189.916	188
8	193.903	197
8.5	197.772	206
9	201.531	208
9.5	205.186	203
10	208.743	203
10.5	212.206	232
11	215.580	217
11.51	218.934	222
12	222.078	228
12.5	225.210	252
13	228.268	237
13.5	231.255	244
14	234.173	237
14.5	237.027	244
15	239.817	240
15.5	242.547	228
16	245.218	246
16.5	247.833	258

NORTHWEST-SOUTHEAST DIRECTION

h	theoretical gamma(nw-se)	actual e-135
0		
0.18	100.261	46
0.56	102.336	115
1.02	107.521	142
1.51	115.984	131
2	127.020	140
2.51	140.586	143
3	154.894	157
3.5	169.990	156
4	184.795	187
4.5	198.565	182
5	210.653	215
5.49	220.382	233
6	227.937	233
6.5	232.640	238
7	234.696	248
7.5	234.305	250
8	231.818	235
8.5	227.698	217
9	222.485	233
9.5	216.751	228
10	211.056	218
10.5	205.898	189
11	201.686	227
11.51	198.665	222
12	197.135	217
12.5	196.987	237
13	198.169	221
13.5	200.471	226
14	203.601	217
14.5	207.214	211
15	210.944	205
15.5	214.439	209
16	217.390	210
16.5	219.555	204

17	250.393	275	17	220.782	213
17.5	252.900	254	17.5	221.009	220
18	255.357	254	18	220.272	218
18.5	257.764	258	18.5	218.692	227
19	260.123	260	19	216.464	219

COORDINATES h_u' and h_v'

EAST-WEST DIRECTION

h_u	h_v	h_u'	h_v'
0.000	0	0	0
0.180	0	0.117	0.117
0.560	0	0.364	0.364
1.020	0	0.663	0.663
1.510	0	0.982	0.982
2.000	0	1.300	1.300
2.510	0	1.632	1.632
3.000	0	1.950	1.950
3.500	0	2.275	2.275
4.000	0	2.600	2.600
4.500	0	2.925	2.925
5.000	0	3.250	3.250
5.490	0	3.569	3.569
6.000	0	3.900	3.900
6.500	0	4.225	4.225
7.000	0	4.550	4.550
7.500	0	4.875	4.875
8.000	0	5.200	5.200
8.500	0	5.525	5.525
9.000	0	5.850	5.850
9.500	0	6.175	6.175
10.000	0	6.500	6.500
10.500	0	6.825	6.825
11.000	0	7.150	7.150
11.510	0	7.482	7.482
12.000	0	7.800	7.800
12.500	0	8.125	8.125
13.000	0	8.450	8.450
13.500	0	8.775	8.775
14.000	0	9.100	9.100
14.500	0	9.425	9.425

NORTH-SOUTH DIRECTION

h_u	h_v	h_u'	h_v'
0	0.000	0.000	0.000
0	0.180	0.117	0.063
0	0.560	0.364	0.196
0	1.020	0.663	0.357
0	1.510	0.982	0.529
0	2.000	1.300	0.700
0	2.510	1.632	0.879
0	3.000	1.950	1.050
0	3.500	2.275	1.225
0	4.000	2.600	1.400
0	4.500	2.925	1.575
0	5.000	3.250	1.750
0	5.490	3.569	1.922
0	6.000	3.900	2.100
0	6.500	4.225	2.275
0	7.000	4.550	2.450
0	7.500	4.875	2.625
0	8.000	5.200	2.800
0	8.500	5.525	2.975
0	9.000	5.850	3.150
0	9.500	6.175	3.325
0	10.000	6.500	3.500
0	10.500	6.825	3.675
0	11.000	7.150	3.850
0	11.510	7.482	4.029
0	12.000	7.800	4.200
0	12.500	8.125	4.375
0	13.000	8.450	4.550
0	13.500	8.775	4.725
0	14.000	9.100	4.900
0	14.500	9.425	5.075

15.000	0	9.750	9.750	0	15.000	9.750	5.250
15.500	0	10.075	10.075	0	15.500	10.075	5.425
16.000	0	10.400	10.400	0	16.000	10.400	5.600
16.500	0	10.725	10.725	0	16.500	10.725	5.775
17.000	0	11.050	11.050	0	17.000	11.050	5.950
17.500	0	11.375	11.375	0	17.500	11.375	6.125
18.000	0	11.700	11.700	0	18.000	11.700	6.300
18.500	0	12.025	12.025	0	18.500	12.025	6.475

NORTHEAST-SOUTHWEST DIRECTION NORTHWEST-SOUTHEAST DIRECTION

hu	hv	hu'	hv'	hu	hv	hu'	hv'
0	0	0	0	0	0	0	0
0.127	0.127	0.165	0.127	-0.127	0.127	0	-0.038
0.396	0.396	0.515	0.396	-0.396	0.396	0	-0.119
0.721	0.721	0.938	0.721	-0.721	0.721	0	-0.216
1.068	1.068	1.388	1.068	-1.068	1.068	0	-0.320
1.414	1.414	1.838	1.414	-1.414	1.414	0	-0.424
1.775	1.775	2.307	1.775	-1.775	1.775	0	-0.532
2.121	2.121	2.758	2.121	-2.121	2.121	0	-0.636
2.475	2.475	3.217	2.475	-2.475	2.475	0	-0.742
2.828	2.828	3.677	2.828	-2.828	2.828	0	-0.849
3.182	3.182	4.137	3.182	-3.182	3.182	0	-0.955
3.536	3.536	4.596	3.536	-3.536	3.536	0	-1.061
3.882	3.882	5.047	3.882	-3.882	3.882	0	-1.165
4.243	4.243	5.515	4.243	-4.243	4.243	0	-1.273
4.596	4.596	5.975	4.596	-4.596	4.596	0	-1.379
4.950	4.950	6.435	4.950	-4.950	4.950	0	-1.485
5.303	5.303	6.894	5.303	-5.303	5.303	0	-1.591
5.657	5.657	7.354	5.657	-5.657	5.657	0	-1.697
6.010	6.010	7.814	6.010	-6.010	6.010	0	-1.803
6.364	6.364	8.273	6.364	-6.364	6.364	0	-1.909
6.718	6.718	8.733	6.718	-6.718	6.718	0	-2.015
7.071	7.071	9.192	7.071	-7.071	7.071	0	-2.121
7.425	7.425	9.652	7.425	-7.425	7.425	0	-2.227
7.778	7.778	10.112	7.778	-7.778	7.778	0	-2.333
8.139	8.139	10.580	8.139	-8.139	8.139	0	-2.442
8.485	8.485	11.031	8.485	-8.485	8.485	0	-2.546
8.839	8.839	11.490	8.839	-8.839	8.839	0	-2.652
9.192	9.192	11.950	9.192	-9.192	9.192	0	-2.758
9.546	9.546	12.410	9.546	-9.546	9.546	0	-2.864
9.899	9.899	12.869	9.899	-9.899	9.899	0	-2.970

10.253	10.253	13.329	10.253	-10.253	10.253	0	-3.076
10.607	10.607	13.789	10.607	-10.607	10.607	0	-3.182
10.960	10.960	14.248	10.960	-10.960	10.960	0	-3.288
11.314	11.314	14.708	11.314	-11.314	11.314	0	-3.394
11.667	11.667	15.167	11.667	-11.667	11.667	0	-3.500
12.021	12.021	15.627	12.021	-12.021	12.021	0	-3.606
12.374	12.374	16.087	12.374	-12.374	12.374	0	-3.712
12.728	12.728	16.546	12.728	-12.728	12.728	0	-3.818
13.081	13.081	17.006	13.081	-13.081	13.081	0	-3.924
13.435	13.435	17.466	13.435	-13.435	13.435	0	-4.031

ISOTROPIC SEMIVARIOGRAMS:

NOTE: These are the transformed isotropic semivariograms corresponding to the original anisotropic experimental semivariograms with their directions in brackets.

h	gamma(e-w)	gamma(n-s)	gamma(ne-sw)	gamma(nw-se)
0	0	0	0	0
0.136	103.361	105.958	104.789	100.016
0.424	110.456	114.568	112.242	100.120
0.772	119.045	123.205	119.822	100.370
1.144	128.190	131.329	126.924	100.788
1.515	137.323	138.731	133.321	101.360
1.901	146.802	145.869	139.400	102.117
2.272	155.858	152.296	144.780	102.999
2.651	165.010	158.493	149.876	104.053
3.029	174.022	164.377	154.624	105.258
3.408	182.823	169.987	159.065	106.613
3.787	191.326	175.355	163.231	108.113
4.158	199.263	180.404	167.071	109.720
4.544	206.970	185.456	170.838	111.533
4.923	213.815	190.226	174.321	113.445
5.301	219.763	194.828	177.614	115.484
5.680	224.592	199.274	180.730	117.645
6.059	228.045	203.575	183.683	119.923
6.437	229.827	207.741	186.485	122.312
6.816	229.608	211.779	189.145	125.524
7.195	230.000	215.697	191.673	128.100
7.573	230.000	219.501	194.077	130.769

7.952	230.000	223.197	196.365	133.521
8.331	230.000	226.792	198.544	136.352
8.717	230.000	230.358	200.662	139.312
9.088	230.000	233.693	202.602	142.218
9.467	230.000	237.009	204.491	145.238
9.845	230.000	240.239	206.294	148.306
10.224	230.000	243.389	208.016	151.415
10.603	230.000	246.461	209.662	154.556
10.981	230.000	249.458	211.234	157.723
11.360	230.000	252.384	212.738	160.906
11.739	230.000	255.240	214.176	164.100
12.117	230.000	258.030	215.552	167.295
12.496	230.000	260.755	216.869	170.485
12.875	230.000	263.419	218.131	173.661
13.253	230.000	266.022	219.339	176.817
13.632	230.000	268.568	220.496	179.945
14.011	230.000	271.058	221.605	183.038
14.389	230.000	273.495	222.668	186.089

SECTION 28 APPENDIX D: DATA FOR THE KRIGING AND AUXILIARY FUNCTION EXAMPLE

EASTING	NORTHING	ln(IP)	IP
492.73	4289.94	3.045	21.010
479.58	4317.46	3.219	25.003
476.26	4310.35	3.638	38.016
473.22	4313.38	3.664	39.017
461.32	4310.62	3.989	54.001
486.75	4306.13	4.060	57.974
473.68	4314.52	4.094	59.979
471.49	4309.44	4.094	59.979
470.27	4310.18	4.111	61.008
473.23	4316.06	4.317	74.963
485.85	4285.48	4.605	99.983
484.56	4295.56	4.605	99.983
473.03	4321.95	4.779	118.985
471.91	4325.03	4.779	118.985
495.49	4295.52	4.787	119.941
465.19	4331.5	4.942	140.050
458.63	4280.92	4.942	140.050
472.18	4320.84	4.942	140.050
479.28	4313.98	4.942	140.050
462.10	4282.72	5.004	149.008
471.64	4317.51	5.011	150.055
464.69	4290.66	5.063	158.064
459.35	4279.19	5.063	158.064
472.69	4307.89	5.075	159.972
465.40	4309.69	5.130	169.017
465.82	4284.28	5.193	180.008
473.58	4321.76	5.193	180.008
480.87	4314.34	5.231	186.980
482.99	4293.19	5.242	189.048
465.83	4280.85	5.298	199.937
480.00	4313.67	5.298	199.937
466.37	4281.04	5.298	199.937
466.37	4282.52	5.298	199.937
471.08	4322.35	5.298	199.937
488.07	4282.64	5.298	199.937
459.92	4282.24	5.347	209.977

484.38	4298.03	5.352	211.030
465.09	4276.98	5.468	236.986
481.33	4313.66	5.468	236.986
468.60	4283.77	5.472	237.936
484.46	4291.4	5.476	238.889
483.00	4296.27	5.476	238.889
472.86	4308.32	5.561	260.083
476.06	4314.6	5.561	260.083
464.72	4280.03	5.561	260.083
484.06	4271.43	5.609	272.871
472.05	4324.63	5.613	273.965
458.18	4282.46	5.635	280.059
473.51	4322.56	5.635	280.059
486.97	4292.57	5.635	280.059
457.88	4280.37	5.635	280.059
471.89	4272.91	5.697	297.972
480.33	4310.58	5.704	300.065
466.10	4281.56	5.704	300.065
477.60	4280.26	5.771	320.858
458.76	4281.97	5.784	325.057
487.72	4293.62	5.784	325.057
480.68	4283.95	5.814	334.956
456.73	4277.82	5.829	340.019
487.63	4279.81	5.835	342.065
463.54	4285.52	5.886	359.963
488.27	4292.41	5.914	370.184
485.02	4291.74	5.919	372.040
480.50	4311.63	5.927	375.028
496.38	4286.82	5.991	399.814
495.78	4294.78	5.991	399.814
476.62	4317.19	5.991	399.814
478.45	4319.34	5.991	399.814
473.06	4274.14	6.009	407.076
484.27	4293.62	6.021	411.990
486.24	4289.92	6.021	411.990
484.97	4292.97	6.033	416.964
473.80	4278.24	6.035	417.799
484.77	4288.69	6.043	421.155
466.43	4329.83	6.064	430.092
485.80	4288.29	6.064	430.092
484.16	4296.46	6.091	441.863
481.72	4295.85	6.114	452.144
459.65	4276.63	6.118	453.956
482.03	4294.18	6.133	460.817
481.83	4289.46	6.136	462.201
485.30	4288.29	6.161	473.902
487.78	4301.32	6.186	485.899
478.20	4304.85	6.192	488.823
483.77	4295.41	6.194	489.801

459.46	4281.19	6.213	499.197
476.70	4319.72	6.215	500.196
485.56	4285.48	6.215	500.196
463.87	4285.02	6.215	500.196
474.78	4320.62	6.215	500.196
464.88	4284.22	6.215	500.196
461.14	4278.56	6.275	531.126
464.83	4284.1	6.286	537.001
460.79	4282.88	6.328	560.035
482.73	4294.42	6.332	562.280
485.76	4287.24	6.365	581.145
483.52	4294.7	6.390	595.857
464.24	4323.15	6.397	600.042
475.29	4323.45	6.397	600.042
473.45	4303.67	6.593	729.968
487.92	4283.63	6.615	746.205
465.23	4287.7	6.620	749.945
489.01	4287.23	6.620	749.945
484.35	4297.38	6.620	749.945
476.35	4314.79	6.620	749.945
482.05	4313.69	6.633	759.758
486.44	4292.17	6.771	872.184
485.63	4300.06	6.838	932.622
485.19	4284.09	6.867	960.064
477.67	4304.64	6.892	984.368
479.58	4308.95	6.892	984.368
467.95	4284.08	7.244	1399.682
487.39	4285.91	7.283	1455.348
473.39	4292.88	7.313	1499.669
479.29	4299.09	7.313	1499.669

SECTION 29 BIBLIOGRAPHY

1. Noel A.C. Cressie, "Statistics for Spatial Data," John Wiley & Sons, (1993).

2. Jonathan D. Cryer, "Time Series Analysis," Duxbury Press, Boston, (1986).

3. Clayton Deutsch and A. G. Journel, "Geostatistical Software Library and User's Guide," Oxford University Press, (1992).

4. Michael Edward Hohn, "Geostatistics and Petroleum Geology," Van Nostrand Reinhold, New York, (1988).

5. A.G. Journel and C.J. Huijbregts, "Mining Geostatistics," Academic Press, London, (1978).

6. W. Magnus, F. Oberhettinger, and F. G. Tricomi, "Tables of Integral Transforms," McGraw-Hill, New York, (1954).

7. J.M. Rendu, "An Introduction to Geostatistical Methods of Mineral Evaluation," South African Institute of Mining and Metallurgy Monograph Series, Geostatistics 2 (1981).